

THESIS



This is to certify that the
thesis entitled

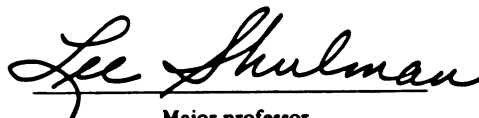
AN OBSERVATIONAL STUDY OF THE RELATIONSHIP
BETWEEN DIAGNOSIS AND REMEDIATION IN READING

presented by

Annette B. Weinshank

has been accepted towards fulfillment
of the requirements for

DOCTOR OF PHILOSOPHY degree in EDUCATIONAL PSYCHOLOGY


Major professor

Date November 19, 1979



JUL 17 9 2006

OVERDUE FINES:

25¢ per day per item

RETURNING LIBRARY MATERIALS:

Place in book return to remove
charge from circulation record

AN OBSERVATIONAL STUDY OF THE RELATIONSHIP
BETWEEN DIAGNOSIS AND REMEDIATION
IN READING

By

Annette Barshefsky Weinshank

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree

DOCTOR OF PHILOSOPHY

Department of Educational Psychology

1980

ABSTRACT

AN OBSERVATIONAL STUDY OF THE RELATIONSHIP BETWEEN DIAGNOSIS AND REMEDIATION IN READING

By

Annette Barshefsky Weinshank

Purpose

This study empirically investigated the clinical problem solving behavior of reading specialists during a total of twenty-four observational sessions as they (1) diagnosed simulated cases of reading difficulty, (2) prepared an initial remediation plan, and (3) associated given remedial statements with diagnostic statements. The central purpose of the study was to assess whether reading specialists' diagnoses lead directly to remedial recommendations within and across sessions for given simulated cases.

Procedures

Eight experienced, credentialled, practicing reading specialists participated in the study. Four received their training entirely in Michigan, and four entirely in Illinois. Each clinician performed a series of tests using three simulated cases of reading difficulty over a period of at least one week. Two of the cases were thinly disguised versions of the same reading problem. The third represented a different reading problem. During each of her three sessions, each subject was asked to:

1. reach a judgment about a case and write a diagnosis;
2. write an initial remediation plan;
3. number and code all key diagnostic statements and transfer them to a standardized diagnostic checklist;
4. number and code all key remedial statements and transfer them to a standardized remedial checklist;
5. give the number and code of the diagnostic statement or statements (if any) associated with a given remedial statement;
6. respond to questions dealing with associations and non-associations between remedial and diagnostic statements; and
7. comment on three free-response questions.

Major Findings

1. A small proportion of the total diagnostic and remedial statements/associations made accounted for agreement across two or more of the six sessions for a given case. The bulk of the statements/associations for a case were idiosyncratic, i.e., the statements/associations were made in only one session.
2. Across all cases a relatively small number of categories of diagnostic and remedial statements accounted for all statements made in two or more sessions.
3. Examination of common case information lead neither to common diagnoses, common remediations, nor common associations between remediation and diagnosis. Agreement between

and within clinicians ranged from very little to none whatever.

4. Only by aggregating diagnostic and remedial statements/ associations across the six sessions could the outlines of a meager consensus on each case be demonstrated.
5. At the individual clinician level, there was essentially no correlation between diagnosis and remediation.
6. At the group level, diagnosis and remediation showed a modest level of association.
7. Clinicians never followed their stated plans regarding information collection procedures and the writing of the diagnosis and remediation.
8. There was no difference in performance between the Michigan and Illinois subjects.

© Copyright by

ANNETTE BARSHEFSKY WEINSHANK

1980

Ben Zoma said, "Who is wise? He who learns
from all men; as it is said, from all my teachers
I have gotten understanding."

Avoth: Sayings of the Fathers, IV, 1.

To teachers from whom I continue to learn,
Professors

Lee Shulman

John Vinsonhaler

Arthur Elstein

George Sherman,

this dissertation is dedicated.

ACKNOWLEDGEMENTS

I am indebted to a diverse band of colleagues and associates. The members of my doctoral committee, Professors Lee Shulman, John Vinsonhaler, Arthur Elstein, and George Sherman are exceptional. They delight in the work of scholar and teacher and have shared both roles freely. My graduate years have been a joy and delight thanks, in large measure, to their example and company.

The computer systems staff of the Clinical Studies Group of the Institute for Research on Teaching at Michigan State University, Chris Wagner and Francine Kitchen, were always ready to create or modify the computer programs necessary for my analysis of the data presented in this study. Ruth Polin, office manager for the group, and Debbie Salters, secretary, were always game for one more editing of computer printout or yet another round of data entry.

The final thesis typing, that long-awaited event, was efficiently and competently done by Linda Raab.

This dissertation reports the results of an observational study, which means that people had to agree to act as subjects and be observed in the performance of certain tasks. It was my great good fortune to be able to work with eight reading specialists who were not only highly professional and competent practitioners, but who also were gracious

enough to express interest in, and encouragement of, the work being done in the study.

A key requirement in the design of this investigation was the replication of the study using subjects who had received all their graduate training outside of Michigan. I was able to carry out this phase of the dissertation thanks to the generous help and cooperation of Professor Ahmed Fareed, chairperson of the Department of Reading, Northeastern Illinois University, Chicago, Illinois. Professor Fareed extended every courtesy to me during my data-collection weeks at Northeastern in May and June of 1979, making my stay both a profitable and pleasurable one. A very special thanks are in order, too, for members of his staff: Dr. Wayne Berridge, who let me take over his office for a number of observational sessions, and Ms. Dorothy Bacon, whose superb organizational skills contributed so much to the success of my visit.

Finally, I salute my family. My sons, Joshua and Joel, provided a fine corrective to the rigors of my doctoral program by remaining steadfastly more interested in their lives than in mine these past years. To my husband, Don--teacher, scholar, and boon companion--I offer singular thanks for always sharing, but never appropriating, my times of doubt and worry and my moments of success and excitement.

TABLE OF CONTENTS

	Page
LIST OF TABLES	viii
LIST OF FIGURES	x
CHAPTER ONE: THE PROBLEM	1
Introduction	1
Background of the Study	3
Overview of the Study	5
CHAPTER TWO: AN HISTORICAL PERSPECTIVE ON THE ASSOCIATION BETWEEN DIAGNOSIS AND REMEDIATION IN READING: 1900-1979	6
Introduction	6
1890-1910	7
1910-1925	7
1925-1935	10
1935-1950	11
1950-1965	13
1965-1979	15
Background of the Current Study	25
CHAPTER THREE: DESIGN AND METHOD	30
Introduction	30
Research Setting	32
Design of the Study	33
The Subjects	34
Method	35
Data Analysis	40
Agreement Statistics	42
Additional Analytic Procedures	45
Summary	45
CHAPTER FOUR: RESULTS	47
Introduction	47
Proportional Agreement	55
(1) Diagnosis	55
(2) Remediation	57
(3) Cues	63
(4) Remediated Diagnosis	66
(5) Specific Remedial/Diagnostic Associations	68

	Page
Characteristics of Unassociated Remedial and Diagnostic Statements	76
Characteristics of Associated Remedial and Diagnostic Statements	80
Commonality	81
Intercorrelations	82
Comparison of Michigan and Illinois Subjects on All Products	89
Correlational Data	91
(1) By Individual	91
(2) By Group	92
Clinician Self-Reports	92
Comments About the Sessions	98
 CHAPTER FIVE: DISCUSSION	 99
 CHAPTER SIX: SUMMARY AND RECOMMENDATIONS	 105
Introduction	105
Summary of Procedures	105
Summary of Results	106
Proportional Agreement	106
Commonality	107
Intercorrelations	108
Comparison of Subjects	108
Correlational Data: Individual and Group	108
Clinician Self-Reports	109
Conclusions	109
Recommendations	111
 APPENDICES	 115
A. CASE INFORMATION INVENTORY	115
B. OBSERVATIONAL STUDY INSTRUCTIONS	118
C. CUE COLLECTION FORM	131
D. SAMPLE DIAGNOSIS	132
E. DIAGNOSIS CHECKLIST	133
F. SAMPLE REMEDIATION	137
G. REMEDIAL CHECKLIST	138
H. REMEDIATION/DIAGNOSIS ELEMENT FORM: SESSIONS ONE AND TWO ONLY	 145

	Page
I. DIAGNOSTIC/REMEDIAL ELEMENT FORM: SESSION THREE ONLY	146
J. DIAGNOSIS ONLY/REMEDICATION ONLY ELEMENT FORM: SESSIONS ONE AND TWO	147
K. REMEDIATION AND DIAGNOSIS GUIDE SHEET	148
L. REMEDIATION ONLY GUIDE SHEET AND RESPONSE FORM . . .	149
M. DIAGNOSIS ONLY GUIDE SHEET AND RESPONSE FORM	151
N. CALCULATION OF PHI CORRELATION	153
O. MASTER DIAGNOSIS AND REMEDIATION	156
P. PROPORTIONAL AGREEMENT FIGURES	158
Q. INTERCORRELATION RANGES	164
BIBLIOGRAPHY	165

LIST OF TABLES

Table	Page
1. Overall Design of the Study	34
2. Summary of Contents of Sessions	34
3. Profile of Participating Subjects	36
4. Proportion of Diagnostic Statements Mentioned in Two or More Sessions	55
5. Proportion of Weakness Statements in Set of Diagnostic Statements	56
6. Agreed-Upon Diagnostic Statements Across Cases	58
7. Proportion of Remedial Statements Mentioned in Two or More Sessions	57
8. Proportion of Treatment Statements in Set of Remedial Statements	59
9. Agreed-Upon Remedial Statements Across Cases	61
10. Proportion of "Other/OTHER" Statements in Set of Checklist Statements	62
11. Proportion of Cues Mentioned in Two or More Sessions	64
12. Proportion of Requests for Test Booklets/Audio Tapes .	64
13. Proportion of Remediated Diagnoses to Total Diagnostic Statements	67
14. Proportion of Remediated Diagnoses Mentioned in Two or More Sessions	67
15. Proportion of Remedial/Diagnostic Associations (Two or More Sessions)	68
16. Proportion of Treatment/Weakness Associations in Set of All Remedial/Diagnostic Associations	69

Table	Page
17. Proportions of Idiosyncratic and Agreed-Upon Statements Across Products	73
18. Summary Table: Proportion of Statements for Each Session Across Products	74
19. Total Number of Statements for Each Session Across Products	75
20. Unmatched Diagnostic Statements	77
21. Mean Commonality Scores for All Products	83
22. Mean Inter-Clinician Correlation	85
23. Partial Intercorrelation Matrix for Diagnosis	86
24. Mean Intra-Clinician Correlation	87
25. Partial Intracorrelation Matrix for Diagnosis	88

LIST OF FIGURES

FIGURE	PAGE
1. Proportional Agreement: All Products	72

CHAPTER ONE

THE PROBLEM

Introduction

This dissertation was designed to answer one question: is remediation reliably related to diagnosis in reading? If the two were found to be closely related then the almost constant assertion in the literature reviewed in this study, namely that diagnosis is an indispensable prerequisite to remediation, would be demonstrated empirically. If, on the other hand, diagnosis and remediation were found to be unrelated, or minimally related, as has been suggested by a few contemporary researchers (Chapter Two) then it would be necessary to (1) determine in what ways the two activities were unrelated, (2) account for the findings, and (3) examine the implications of the findings for the continuation of current practices in the areas of both clinical care of children with reading difficulties, and of teacher education in reading.

Since its origin within the European medical community early this century (Chapter Two), the literature on remedial reading has persistently endorsed the necessity of obtaining two (dia) types of knowledge (gnosis) as a prerequisite for effective intervention: knowledge of symptoms and knowledge of underlying causes (Stein, 1973). Accordingly, even at present, training procedures for reading specialists continue to stress the necessity of gathering information from a variety of

sources including information from schools, parents, agencies working with the family, health and developmental history, and analytic testing in a variety of areas including personality (CEMREL, 1979). Considerable resources of time and money are spent in reading clinics and schools for specialists to gather information about children, administer tests to them, and prepare diagnoses and suggested remediations.

The major justification for all this activity would appear to be the stated conviction that effective remediation can proceed only after a thorough diagnosis has been prepared. The fundamental assumption, then, is that diagnosis and remediation are so strongly related that to proceed to treat in the absence of an eclectic and thorough diagnosis would be both ineffective and irresponsible.

What if this assumption is incorrect?

If diagnosis and remediation as presently executed are shown to be unrelated, serious issues would be raised about the usefulness of the approach taken by present programs devoted to training in the diagnosis and remediation of reading difficulties. The value of extended diagnostic testing sessions and data gathering activities in clinics would come under question. The role of the reading consultant in the schools would have to be reevaluated. In sum, demonstrating the inaccuracy of the assumption would provoke marked consequences throughout the educational system.

This study empirically investigated the clinical problem solving behavior of reading specialists as they examined information contained in simulated cases (SIMCASES) of reading difficulty in order to demonstrate whether or not any reliable relationship existed between

diagnosis and remediation. Specifically, five questions were seen to be embedded within the major question posed:

1. How reliable, or consistent were the diagnoses?
2. How reliable were the remediations?
3. How reliable were information collection procedures?
4. How reliable were diagnostic statements to which remediations had been attached?
5. How reliable were specific remedial/diagnostic associations?

Additionally, three related questions were to be addressed:

1. What was the strength of correlation between diagnostic and remedial performance?
2. Was there a discernable institutional training effect on performance?
3. How did the specialists themselves view the tasks of diagnosis and remediation as they engaged in those activities?

To obtain the necessary data, eight experienced, credentialed, practicing reading specialists, four from Michigan and four from Illinois, were asked to examine three simulated cases of reading difficulty in order to (1) make diagnostic judgments, (2) prepare initial remediation plans, and (3) associate given remediations with diagnostic statements. In all, 24 sessions were conducted.

Background of the Study

Studies conducted over the past four years by the Clinical Studies Group of the Institute for Research on Teaching have developed techniques for studying and measuring the diagnostic clinical problem solving behavior of (1) reading specialists (Vinsonhaler, 1979;

Weinshank, 1978; Hoffmeyer, 1979), (2) learning disabilities specialists (Van Roekel, 1979), and (3) classroom teachers (Gil, 1979; Stratoudakis, 1979). These investigations were based on the Inquiry Project studies in medical problem solving conducted at Michigan State University (Elstein, Shulman, and Sprafka, 1978; Vinsonhaler, Wagner, and Molitor, 1975).

The results of the medical studies led to the formulation of the Inquiry Theory of clinical problem solving (Vinsonhaler, Wagner, and Elstein, 1978). The Inquiry Theory proposed that clinical diagnosis is determined by the interaction of (1) clinical memory (consisting of sets of problems, cues, cue values, diagnoses, treatment and the relationships among them), (2) clinical strategy (the sequencing of the mental tasks performed by the clinician), and (3) the case (Gil, Hoffmeyer, Van Roekel, and Weinshank, p. 7, 8).

According to the theory, adequate knowledge of the relationship between cues (case information) and problems is central to correct diagnosis, and adequate knowledge of the relationship between the same problems and their treatments is central to effective remediation. In other words, the case information gathered by the clinician should suggest likely problems and these, in turn, should be associated in memory with effective treatments. In short, according to this formulation, treatments are dependent upon the identification of problems. Put another way, remediation is dependent on prior diagnosis.

The central question of this dissertation was whether or not reading specialists' diagnoses did, in fact, lead directly and reliably to remedial recommendations as the Inquiry Theory predicts they should, and as the literature contends they must.

Overview of the Study

In Chapter Two, the growth of the field of remedial reading is traced from its origins early this century within the European medical community to the present. The variations on the still dominant diagnosis-treatment model are examined. The chapter concludes with a report of a pilot study (Weinshank, 1979) the results of which call into question the basic assumptions of the diagnosis-remediation model and establish the guidelines for the present study.

In Chapter Three, the research questions posed by the dissertation are reviewed, The research setting in which the data were collected is described, together with the design and the methods used for data gathering and analysis.

The results reported in Chapter Four speak directly to the sequence of questions posed at the beginning of Chapter One.

Speculations on some alternative explanations for the study's results are discussed in Chapter Five. Methodological limitations are also discussed.

A summary of the study and recommendations for further work are presented in Chapter Six.

CHAPTER TWO

AN HISTORICAL PERSPECTIVE ON THE ASSOCIATION BETWEEN DIAGNOSIS AND REMEDIATION IN READING: 1900-1979

Introduction

The idea that children having difficulty in reading must first be diagnosed before being given special teaching is a twentieth century phenomenon in the United States. The field of remedial reading as reflected in the literature has had at its base the stated belief that diagnosis and remediation are bound together. Deficiencies within the learner can be detected, their causes ascertained, and corrective action taken in order that a more desirable level of proficiency results. The diagnosis-treatment model has directed both thought and practice in the field since its origin early this century within the European medical profession, when congenital alexia, or 'word blindness' was described as being the cause of retardation in reading (Nila Smith, 1967). Remediation consisted of training children in a technique called the alphabetic-spelling method.

The diagnosis-treatment model has undergone a number of permutations during the course of this century but the basic assumption underlying the model has remained intact: diagnosis (whatever the scope) and treatment (however conducted) must be related.

Nila Smith and Harris (1967, 1977) present a chronology of instructional and diagnostic-remedial eras in the teaching of reading

during the twentieth century. The review of the literature presented in this chapter will be organized according to their chronologies.

1890-1910

The cultural influences during this period mandated that reading serve two main purposes: (a) that it enable the reader to gather knowledge from the author as printed, and (b) that the reader be able to reproduce and impart this knowledge to others in a clear and pleasing manner (N. Smith, Chapter 5). Expressive oral reading was taken as an indicator that the reader was accomplishing both those goals and by so doing was developing literary appreciation. Children taught by the word method during initial reading instruction were judged as not being able to read well in the upper grades primarily because, it was argued, they had not become independent word callers. The result was the adoption by some writers of an extremely synthetic phonics system which began not with the whole word but rather with the sounds of the letters.

The program of instruction was hierarchically arranged from letters to syllables and from phonetically regular to phonetically irregular words.

1910-1925

A startlingly abrupt reversal from oral to almost exclusively silent reading occurred during this period. The new emphasis was on meaning-getting in all phases of education. For the first time, research results were instrumental in changing reading instruction: Laboratory studies indicated that silent reading was superior to oral

reading both in speed and comprehension. Thus, the emphasis on oral reading as the indicator of accurate acquisition of knowledge was swept aside along with the extensive phonics training which was believed necessary for achieving expressive oral reading. Experience charts were introduced and instruction focused on reading sentences, phrases, and words. The overwhelming emphasis was on skillful silent reading for meaning, with comprehension being measured in a variety of ways.

During this period of 1910-1925, the term "remedial reading" first appeared. A paper by Uhl (1916) was apparently the first dealing with the idea that children having difficulty in reading could be diagnosed and given special teaching in a school setting.

Uhl reported that each student received individual help 15 minutes daily for two to three weeks, with treatment "continued or modified during the training period according to the apparent needs of the pupils...While the number of pupils in each of the grades is small, yet the presence of a large relative gain in each grade upon the part of those drilled seems significant."

Uhl drew a number of conclusions from the results of the study: (1) an accurate diagnosis is such that the specific defects of poor readers can be detected; (2) carefully planned individual treatment will produce as rapid growth as is produced in the case of apparently brighter students by class instruction; and (3) drill should be carried on during the school day by those who are experienced in teaching the subject.

The Twenty-Fourth Yearbook of the National Society for the Study of Education (Whipple, 1925) reflected the concern with reading

deficiency uncovered by the Army's tests of recruits during World War I. The Yearbook firmly endorsed the position that, "Silent reading is of largest social value...It is evident that there is need of vigorous emphasis in any program of instruction of habits of intelligent silent reading." Nevertheless, the authors cautioned that in beginning reading instruction and in diagnostic and remedial cases, certain "habits" must be given specific attention. These include recognition of sentences as units of thought, recognition of words and groups of words, and recognition and interpretation of typographical devices such as punctuation, paragraphing, etc.

The objective of remedial work in the opinion of the Yearbook's authors was the removal of a deficiency by first analyzing the particular causes of that deficiency. Their aim was to encourage teachers to make a systematic attack on the problem cases in their own classes. Individual examination via standardized tests, informal tests, and personal interview would hopefully reveal fundamental attitudes and causes of deficiency. Specific remedial measures attacking the cause of deficiency would be formulated and initiated, followed by measurement of progress and finally adjustment of work to changing needs until the deficiency was removed.

In the field generally during this period diagnosis usually involved the compilation of a case history (school, home, medical); the administration of standardized reading tests, and observations of such motor factors as eye-movements, vocalization, extraneous bodily movements, and breathing.

Remedial measures were of three different types: (1) psychologists were using special devices, including the kinesthetic method for teaching phonics; (2) medical people were using the alphabet-spelling method for "dyslexiacs"; and (3) educators concerned with instruction for retarded readers in the classroom were using a variety of procedures to improve oral reading, silent reading, word recognition, and rate. They also emphasized methods designed to remedy inadequate eye movements, extraneous bodily movements, and improper breathing. However, authors were criticized for making only general suggestions for diagnostic problems and not reporting remedial techniques in enough detail to make them useful. The overall conclusion was that remedial methods had not kept pace with diagnostic methods (Harris, 1967).

Nevertheless, the diagnosis-treatment model itself remained undisputed. The question of which, if any, diagnostic findings actually impacted the course of treatment was never raised.

1925-1935

Remedial reading was a chief subject of study during this period. (1) Educational clinics began to be established because it was perceived that the needs of many poor readers could not be uncovered using classroom diagnosis. (2) Batteries of tests for use in diagnosing reading difficulties were developed. These included the Gates Reading Diagnosis Tests and the Durrell Analysis of Reading Difficulty. (3) Theories of organic causes of reading disability emerged including visual, auditory, and lateral dominance factors, heredity, mixed dominance of cerebral hemispheres, inadequate mental ability, and abnormal emotional tendencies (Harris, 1967).

Interestingly, during this period of continued emphasis on meaning and subordination of phonics, several major research reports dealing with severely disabled readers (Chall, 1967) concluded that the one characteristic common to almost all poor readers was word recognition and analysis errors. Thus, although theories of causation differed, the recommended remedial training focused on systematic phonics programs, often augmented by kinesthetic and motor aids.

1935-1950

While the cultural emphasis in reading shifted during this period (reading was now seen as necessary for living effectively in a democracy), the use of varied instructional approaches in reading continued, as did the interest in reading disability.

A major effect of World War II on the field was the (re) discovery that thousands of military recruits could not read well enough to follow even simple printed instructions. Concurrently, a number of investigations revealed reading deficiencies in large numbers of high school and college students.

Diagnostically, the major change according to Harris (1967) was the change from an organic-causation to a multiple-causation theory of reading disability. This pluralistic view of causation placed an emphasis primarily on visual, social, and emotional problems and secondarily on inappropriate teaching methods, neurological difficulties and speech and auditory difficulties.

Developments in remediation included: (1) the use of informal diagnosis with basal readers; (2) establishment of large scale remedial programs in public schools, spreading from elementary schools, to

secondary schools and universities; (3) increase in the number of remedial reading teachers; (4) increase in the number of new books on the prevention and correction of reading difficulties; (5) development of mechanical reading pacers; and (6) experimental support for the view that motivated practice resulted in as much improvement as extensive use of special mechanical devices (Harris, 1967).

As was the case for disabled readers during the silent reading period, research published during this period of limited reintroduction of phonics also reported that backward readers were unable to identify and analyze words (Chall, 1967). In these reported cases training was designed to overcome this specific instructional deficiency.

Robinson (1946) paid little attention to phonics, or decoding, versus meaning issues. She used a team approach to diagnosis and remediation incorporating the findings of social worker, psychiatrist, speech correctionist, otolaryngologist, endocrinologist, and reading specialist. The findings on seriously disabled readers emphasized causal factors relating to maladjusted homes, visual anomalies, and social/emotional problems. Incorrect reading methods were fourth in importance. Other causal factors considered to be of lesser importance were speech/auditory functioning, endocrine problems, neurologic impairment, and general physical condition. Robinson's classic study summarized the treatment and follow-up provided for each student in the study based on what were considered to be probable causes of the disability. Almost all these seriously deficient students were reported to have made gains in social and emotional adjustment.

Progress in reading could be characterized as variable, ranging from no progress, through to excellent progress.

Robinson commented that:

Many of the anomalies in allied fields may be discovered and remedied without appreciable growth in reading. This is because such remediation only prepares the child for learning to read and does not teach him the skill. A direct and vigorous reading program must follow correction of causal factors...In many instances, it is probably that an enthusiastic, capable teacher can motivate a child to learn to read even though he has inhibiting difficulties, although much less time and effort might be needed if the inhibiting factors were corrected prior to learning (236-237).

Robinson's study provided powerful support for the idea of multiple causality, and chiefly social-emotional causality of reading difficulty. Instruction, particularly in severe cases, was seen as necessarily following the removal of deficits in allied areas. While the usefulness of this approach for most children in most classrooms was highly debatable, preference for emphasis on treatment of reading problems by some form of therapy continued for a decade after the study was published.

1950-1965

Smith saw two major influences as shaping reading instruction during this period: Expanding knowledge in all areas of inquiry and the technological revolution. Extraordinary concern about the teaching of reading emerged. Reading was seen as the crucial skill needed to survive in a technologically complex world. No single method of instruction held sway but over-emphasis on phonics continued to be discouraged and emphasis on reading for meaning was reiterated. Also,

systematic instruction in comprehension skills was advocated as well as instruction in reading to suit the requirements of unique content areas.

Interest in reading disability persisted during this period and numerous changes took place. Rudolph Flesch (1955) stated flatly that children couldn't read because they were being taught initially by the whole word method instead of the phonic method. He argued that the prerequisite for mature reading was mastery first of the mechanics of decoding and claimed that all research studies to that time supported that view. Flesch argued, in effect, that mature reading and learning to read were two different activities and that reading instruction was a fiasco because the word method forced children to grasp words as if they were experienced readers. Flesch claimed that strict attention to phonics was the only sensible way to conduct remedial as well as initial reading instruction and devoted the second half of his book to exercises that presented one, two and three syllable words in a carefully graded hierarchy with each new level deriving from sound-symbol associations learned earlier.

Harris believed that Flesch's book was beneficial to the field of remedial reading.

He had convinced hundreds of thousands of parents that when Johnny had trouble with reading he was not necessarily stupid, and this led to public pressure both for improved developmental reading programs and for more diagnostic and remedial facilities (Harris, 1967).

Nevertheless, the vast majority of clinicians and reading specialists continued to subscribe to the multiple-causation theory of reading disability using a variety of diagnostic tests, materials, and methods. Increased effort was made to obtain information related to

reading disability on the part of many investigators specializing in disciplines other than teaching: sociologists, psychologists, and physiologists, etc. Medical researchers studied the effects of hormonal imbalance and the results of administering certain medication to disabled readers, and the congenital word blindness theory re-emerged as one (not the only one as earlier) possible cause of reading disability.

The emphasis on treating reading disability by some form of psychotherapy was gradually replaced by renewed emphasis on physiological, neurological, and constitutional factors. The renewed emphasis on phonics, however, also resulted in an outpouring of new remedial (as well as developmental) programs including the initial teaching alphabet, materials based on the same principles advocated by Flesch, programmed materials, programmed tutoring, and specific phonics systems.

Overall, however, remedial measures again lagged behind diagnostic techniques. Additional training in perception, more attention to vocabulary and phonics and better parental guidance accounted for most of the newer emphases.

1965-1979

Amid the unresolved controversies concerning (1) the efficacy of various modes of instruction, (2) the validity of diverse causal factors, and (3) the long-term effectiveness of various remedial approaches, the Federal Government entered the scene. In 1965 Congress passed the Elementary and Secondary Education Act (ESEA) through which millions of dollars were earmarked specifically for

remedial reading for disadvantaged children. The new cultural influence held that the ability to read provided the key to economic survival and thousands of new positions in remedial reading were created and an immense program of diagnostic and remedial services was inaugurated.

The instructional mode since 1965 could best be described as eclectic: basal, phonics, and language experience systems, individualized programs, trade books using controlled vocabularies, etc. So diverse was the instructional scene that Natchez (1968) edited a large volume specifically excluding discussion of any instructional systems, believing that, "There is no one method with which to deal best with the variety of problems pertaining to reading difficulties, but that effective treatment stems from the teacher's own interpretations of his chosen theoretical framework." The section on causation of reading disorder centered once again on emotional, neurophysiological, and cultural factors alone.

The chapters dealing with diagnostic considerations presented some divergent positions but all agreed that a differential diagnosis must take into account the possibility of emotional, neurophysiological, language, motor, or perceptual disorders as having possible bearing on the reading disability.

In the final section of Natchez' book, the authors dealt with classroom treatment. Without exception they had nothing whatsoever to say about the myriad causal factors and diagnostic considerations presented earlier in the volume. Rather, aside from routine statements about the necessity of attending to the motivation and personal

interests of the child, the authors described techniques or presented rationales for improving word attack and comprehension skills so that (1) reading could become pleasurable, and (2) subject matter areas could be mastered. Even those chapters which dealt with severe disability did little more than recommend some elaboration of basic classroom techniques: (1) use of certain kinesthetic methods; (2) allotment of more time-to-mastery for word analysis and blending skills; and (3) provision of adjunctive services in psychotherapy.

The blatant discrepancy between diagnostic categories and treatment options escaped any mention whatever. The ritual of stating that thorough diagnosis must precede treatment was followed, no matter that even on the face of it the remedial suggestions presented made vanishingly little use of any of the information in the prior diagnosis.

At the same time that Natchez was urging teachers to approach diagnosis and remediation from the point of view of their own theoretical framework, a new theoretic framework for approaching reading instruction was being proposed, a theory derived from the study of the interrelationships between thought and language: psycholinguistics.

Goodman (1969) wrote that:

The reader is viewed as a user of language who processes three kinds of information, grapho-phonetic, syntactic, and semantic, as he reacts to the graphic display on the page. In comparing unexpected responses, the psycholinguistic reading process is revealed.

In practice this meant that if an oral reading error preserved the sense of the message, even if the word was a variant of the actual graphic display on the page, the reader had not really made an error. Reading for meaning was paramount; code-breaking was secondary.

By 1973, F. Smith presented what he called "three radical insights:" (1) Only a small part of the information necessary for reading comprehension comes from the printed page; (2) comprehension must precede the identification of individual words; and (3) reading is not decoding to spoken language. The instructional implications of this view were elaborated by Goodman (in F. Smith, 1973). Children were not to be put through a prescribed sequence of objectives or materials.

The child is already programmed to learn to read. He needs written language that is both interesting and comprehensible, and teachers who understand language-learning and who appreciate his competence as a language-learner.

Holmes (in F. Smith, 1973) put the matter succinctly: "Text can be comprehended only if it is read for meaning in the first place; reading to identify words is both unnecessary and inefficient."

Smith (1978) was disinclined to make any specific recommendations about reading instruction.

The question of the best method for teaching reading has, I think, not been answered because it is not the appropriate question; children do not learn because of reading programs but because teachers succeed in preventing programs from standing in children's way. It is the wisdom and intuition of teachers that must be trusted, provided that teachers have the theoretical foundation to reflect upon the decisions that only they can make.

Despite Smith's (and Natchez') reliance on the wisdom of the properly trained practitioner, and despite the addition of psycholinguistics to the armamentarium of approaches to reading instruction, the number of reading disabled children continued to climb during the period 1965 to the present (Satz, 1976; Guzak, 1978).

Once again in the remedial reading field there continued to be a disinclination, among many writers, to regard quality of instruction as a causal factor. It should be kept in mind that numerous specialists outside the field of reading had become involved in the search for causal factors believed to originate within children and their environment. The field had defined and populated itself in a way that made other causes of reading difficulty (such as quality of instruction) appear as perhaps less fruitful areas of inquiry. Sources of causation of reading disabilities (Harris, 1977) continued to focus on multiple causation including physiological, social, and neurological components, while the learning disabilities movement emphasized deficiencies in one or more of the basic psychological processes involved in understanding or using spoken or written language in children of normal intelligence, physical condition, environment, and emotional make-up.

Satz (1976) wrote, "The incidence of learning disorders is not decreasing, in fact, it may be increasing. Moreover, the effectiveness of current remediation or intervention programs is quite variable, if not discouraging." Satz believed there was an urgent need for early detection and intervention within a longitudinal context that incorporated both short and long-term gains. However, Harris (1977) pointed out that at present it is not yet known if any of the screening procedures used to predict reading failure specifically is more accurate in predicting failure than a good group reading test. He further considered that the number of false negatives (children incorrectly labelled as needing help at a future date) would be unacceptably high.

Yet another set of psychological processes that may cause reading disability was advanced by Downing (1973). He saw the motivation of the literacy learner as being of such importance that arguments could be made for making motivation the chief factor in success and failure in learning to read. "Clearly the place of literacy in a culture's priorities in education and the different ways in which these are implemented are bound to exert different motivational pressures on the learner" (p. 81). He pointed out that in the United States and France referrals of boys far exceeds that of girls but in Germany, Nigeria, and India the reverse is true. Downing attributed these disparities to cultural expectations for younger children.

At the very least, it can be concluded that if there are any innate constitutional differences between girls and boys that effect their development of language and reading skills, they can be outweighed by other factors, as they must have been in countries like Germany, Nigeria, and India (p. 110).

Finally, Downing made reference to the disagreement among experts internationally as to what is meant by the terms "reading" and "literacy." Elgin (1978) maintained that literacy is an attitude toward reading which sees the absence of books as being an unhappy event in one's life. By that criterion, she argued, the United States is an illiterate society. If Downing's stress on motivation is accurate, and the purpose of reading in this country is not for personal enhancement, pleasure, or solace, then this would constitute an important causal influence on the attitude toward reading of students and teachers alike.

The current statistics on reading disability in the United States are of awesome proportions: 15 percent of American schoolchildren

(eight million) are classified as having severe reading problems (Satz, 1976). Twenty percent of the adult population (23 million) are functionally illiterate, unable, for example, to read newspaper help-wanted ads. Less than half of the nation's total adult population aged 18 to 65 (55 million) are really proficient in reading skills (Guszk, 1978).

"The present state of affairs is such that there can be no assurance that a diagnosis will be accurate or that remedial instruction will be sufficient to meet a child's needs" (Satz, 1976).

In a statement that could have been inserted into Uhl's article or the twenty-fourth NSSE Yearbook, Rutherford (1972) wrote:

When a teacher explicates a child's reading problem in terms of reading skills that the child does and does not possess, and types of reading activities that he can and cannot perform, then the teacher has attained the desired diagnostic level--the prescriptive level.

Rutherford did not discount the possibility that other causal factors (psychodynamic, sensory, neurological, etc.) were involved. He stressed, though, that, "Diagnosis of reading problems is useful to the classroom teacher only to the extent that it suggests what he can do to improve the child's reading performance." This reappearance of diagnostic and prescriptive reading instruction represents the most recent attempt to preserve the assumptions of the diagnostic-prescriptive model, namely that diagnosis directly informs the remedial process.

For the most part, the recent texts and periodic literature that have dealt with the diagnostic-prescriptive model have finessed the problem of which approach to instruction might be "best." They have

emphasized instead that, whatever the approach, there will be children who don't progress adequately and that it is the classroom teacher who is responsible for correcting the situation.

For example, Harvey (1974) and Bond (1976) both emphasized the role of the teacher and the teacher's knowledge about how to modify instruction as the key elements in the successful teaching of reading. Guszak (1978) assumed that if the diagnostic reading teacher knew the various reading skills (word recognition, comprehension, fluency), had a command of motivational techniques, and could determine accurately pupil possession of such skills, he or she had arrived at the point of prescribing instruction that would develop new behaviors. Guszak considered diagnosis without specific prescription to be a wasted experience.

Other writers continued to emphasize the role of the teacher and to de-emphasize the prominence of extended diagnosis per se. Rosenshine (1970) and Mallaragno (1974) made very limited use of diagnostic evaluations. They both described programs in which it was assumed that children will improve in reading if given systematic, intensive instruction in skill development in both decoding and comprehension. Similarly, Stauffer, Abrams, and Pikulski (1978) believed that only when a well designed instructional plan was unsuccessful should a more comprehensive diagnostic evaluation be considered.

A more moderate position vis-a-vis the usefulness of diagnosis was found in literature that referred to levels of diagnosis. Bond and Tinker (1967) and Ekwall (1976) believed that diagnosis must be efficient, going only as far as necessary. The remedial teacher studies the diagnostic findings and arranges a learning situation that

will enable the student to progress more rapidly than had been the case in the past.

Dechant (1969) described therapeutic diagnosis as being concerned with the present state of the child in order to determine correctable conditions in the child's reading. On the other hand, he was more skeptical of the usefulness of etiological diagnoses, believing that searches for causes of present difficulty were often useless for purposes of formulating a remedial program.

Carter (1970) saw two major approaches to diagnosis and treatment: Clinical and corrective. Both are matched to treatment but the corrective approach is classroom based and does not take into account causal factors. The clinical approach includes causal factors. This dichotomy is similar to Dechant's but is more sympathetic to the usefulness of uncovering etiological factors in reading disability.

Harris (1977) pointed out that skill development via the diagnostic-prescriptive model has brought with it a variety of management schemes. Objectives are stated, pretests and posttests are developed for each objective, and instructional materials are prepared. A child's assignments are based on pretest performance. If posttest mastery is not demonstrated, additional assignments are given. Harris cautioned that:

The temptation to stress highly specific goals which are easy to test sometimes distorts the program into a great over-emphasis on decoding. Much time is probably wasted in rigidly following the idea of a pretest for every objective. There is a strong temptation to go by the number of correct answers and not inquire how or why the child made his errors, so that diagnostic thinking is at a very low level.

Nevertheless, Harris appeared to be convinced that an effective diagnostic-prescriptive program would result in fewer children requiring outside help. The reading specialist would act primarily as a resource person and helper to the teachers and would also work with the few children who required intensive and expert help.

During this same time of 1965 to the present, however, a few authors questioned the existence of any significant relationship between diagnosis and remediation, whether the causal factors were seen as being neurological, psychological, instructional, etc. Strang (1969) wrote, "Research on the diagnostic process itself is practically non-existent. There is little evidence that the recommended remedial techniques, if used, would be effective, or that techniques not recommended would get equally good results." Spache (in Newman, 1968) contended that there was still widespread lack of integration between the two processes of diagnosis and remediation and stated that, "Numerous reports of remedial work give evidence that the procedures used are not directly related to the detailed diagnostic findings."

Bateman (1971) asked:

Was the diagnosis a necessary and sufficient prerequisite to the remediation? Might other remediation, not derived from the diagnosis, have been equally successful? A child is diagnosed but the remediation is not successful. Was the diagnosis inadequate, or was an error made in deriving the remediation? In any case, it is difficult to establish a valid relationship between diagnosis and remediation.

She concluded:

If the diagnostic-remedial approach is found valid, what then is the relationship between diagnosis and remediation? Is effective remediation geared to the child's strengths, weaknesses, neither or both? Almost all have tacitly agreed that diagnosis and remediation ought to be related, either directly

(teach to the strengths) or inversely (teach to the weakness). Unfortunately, there is as yet no direct evidence to support the efficacy of such a matching procedure.

Robinson's study still remains the only long-term investigation in which diagnosis of causal factors in reading disability (primarily emotional) was associated with specific treatment (primarily psychotherapeutic) and the results monitored over a long period of time. She demonstrated, in my view, that treating people with reading problems by providing professional psychological support resulted chiefly in enabling these people to cope with their lives more adequately but did not necessarily result in their becoming better readers.

Similarly, when 25 years later auditory/visual perceptual deficiencies were seen as causing reading disabilities, children remediated with direct teaching of perceptual skills increased their scores on tests that measured this ability but did not increase their reading scores over those of a control group (Guthrie and Siefert, 1978).

Not only in the field of reading was the tacit agreement that diagnosis and remediation are closely related sometimes questioned. Kendell (1975) wrote that:

In most branches of medicine the value of diagnosis is never questioned. Its importance is self-evident because treatment and prognosis are largely determined by it...Where mental illness is concerned, the situation is rather different. The therapeutic and prognostic implications of psychiatric diagnoses are relatively weak, and the diagnoses themselves are relatively unreliable.

Background of the Current Study

The diagnostic-treatment model in remedial reading has been the dominant one since the beginning of the field. The explanations of

what the two tasks ought to encompass (e.g., multiple causality/treatment in allied fields; instructional causes/skill training, etc.) have varied over the years but the underlying assumption has not: diagnosis is an absolute prerequisite to treatment. There have been researchers who questioned this assumption, but until now no empirical evidence has been available either to support or challenge this central assumption.

A first attempt empirically to examine the extent of the relationship between diagnosis and remediation was a pilot study (Weinshank, 1979) which utilized data gathered during an observational study (Vinsonhaler, J., 1979) in which experienced clinicians were asked to write a diagnosis and remediation after examining three simulated cases (SIMCASES) of reading difficulty over a three week period. Two of the cases were superficially disguised replicates of one another.

The major finding was that there was, on the average, a very slight positive association between problems listed in diagnosing the two essentially identical cases. The average correlation for replicated cases was 0.14 (range of 0.00 to 0.30). This measure of consistency represented the degree of reliability with which a clinician arrived at the same diagnosis for the same case.

The major results of the pilot study were:

1. There was a 0.40 correlation between problems stated in a given diagnosis and treatments proposed in a given remediation. The two acts were considerably more independent than had been expected.
2. A very small number of remediations was used repeatedly across cases seemingly irrespective of prior diagnosis.

3. Clinicians who were more reliable diagnostically also tended to be more reliable in terms of remediation plans.

Closer examination of the written diagnoses from the 1977 observational study revealed that not all diagnostic statements were statements of problems. Diagnoses could be broken down into three sub-categories: (1) problems (or weaknesses), e.g., cannot apply phonetic analysis independently; (2) strengths, e.g., uses context well; and (3) observations, e.g., he's the middle child in the family.

Closer examination of the written remediations revealed that not all remediations were statements of treatments, just as not all diagnoses were statements of problems.

Remediation plans were composed of:

1. Problem (re)statements e.g., his verbal ability indicates he will have difficulty in learning to read.
2. Statements of strength e.g., has high average ability and some basic skills.
3. Observations e.g., he seems extremely interested in sports.
4. Treatment statements e.g., help the student become familiar with cv/vc syllabication rule.
5. Problem-treatment statements e.g., help reversal problem pointing out differences in similar words.

- | | |
|--------------------------------------|---|
| 6. Treatment-prescription statements | e.g., build sight vocabulary by using flashcards, word blanks, beat-the-clock, and by incorporation into a short story. |
| 7. Prescription statements | e.g., use language experience stories. |

The four major remediation categories, then, were (1) problems, (2) strengths, (3) observations, and (4) treatments.

A number of hypotheses were generated by the results of the pilot study:

1. Diagnostic statements of weakness (problems) are repeated more reliably across sessions for a case than are statements of strength and observation.
2. Statements of treatment are more reliably associated with diagnostic problems than with strengths or observations.
3. Diagnostic statements to which remediations are attached are repeated more reliably across sessions for a case than are the unremediated diagnostic statements.

Inconsistency in the association of treatments in the initial remediation plan with problems in the diagnosis was hypothesized to arise from two sources:

1. The clinician does not associate a problem in the diagnosis with any remediation. Possible explanations are that the clinician forgot to include a treatment for the problem and/or the clinician doesn't know of any appropriate treatment and/or treatment is not called for.

2. The clinician does not associate a treatment in the remediation with any problem in the diagnosis. Possible explanations are that the clinician never articulated the problem in the diagnosis and/or always uses this treatment irrespective of diagnosis.

The present study was designed to test the observations and hypotheses generated by the preliminary results of the pilot study and to provide more definitive information about the connections clinicians make between remediation and diagnosis.

CHAPTER THREE

DESIGN AND METHOD

Introduction

The aim of this study was to determine whether there was a reliable relationship between diagnosis and remediation in reading. Preliminary data collected during a pilot study (Chapter Two) suggested that only a modest relationship existed between the two tasks. In order to replicate and account for the observed lack of correspondence, a series of precise questions and hypotheses was formulated for this investigation.

1. How reliable, or consistent, are clinicians' diagnoses? Data from earlier studies conducted by the Clinical Studies Group (Chapter One) showed that the diagnostic reliability of reading clinicians' diagnoses was very low.

Hypothesis: Diagnostic statements of weakness (problems) are repeated more reliably across sessions for a case than are statements of strength and observation.

2. How reliable are their remediations? The pilot study indicated that only a small number of treatments was used repeatedly. Might remediation be somewhat more reliable than diagnosis?
3. How reliable are information collection procedures? That is, did the clinicians tend to choose the same case information repeatedly?

4. How reliable are diagnostic statements to which remediations are attached? That is, even if over-all diagnostic reliability were shown to be low, was increased reliability demonstrated for those particular diagnostic statements which were labeled as requiring remediation?

Hypothesis: Diagnostic statements to which remediations are attached are repeated more reliably across sessions for a case than are the unremediated diagnostic statements.

5. How reliable are specific remedial diagnostic associations? In other words, are similar diagnoses paired repeatedly with similar treatments?

Hypothesis: Statements of treatment are more reliably associated with diagnostic problems than with strengths or observations.

6. What is the correlation between diagnostic and remedial performance? The pilot study indicated that those clinicians who exhibited relatively higher reliability in their diagnostic performance also tended to have higher remedial reliability.
7. Does institution attended effect performance? Do clinicians trained in institutions widely separated geographically perform the task of diagnosing and remediating reading difficulties in significantly different ways?
8. How do the reading specialists themselves view the tasks of diagnosis and remediation? Does their performance on the tasks presented to them during the study conform to their own descriptions of what they do?

Specifically, the procedures were as follows:

1. Categorization and standardization of diagnostic statements were obtained by having the clinicians transfer their written diagnoses to a standardized diagnostic check list.
2. Categorization and standardization of remedial statements were obtained by having the clinicians transfer their written remediations to a standardized remedial check list.
3. Associations between remedial and diagnostic statements were obtained from the clinicians.
4. Specific information was obtained about:
 - (a) matched diagnostic-remedial statements;
 - (b) unmatched remedial statements; and
 - (c) unmatched diagnostic statements.

Research Setting

The observational sessions were conducted in a small room containing chairs and a table. The session involved the experimenter and the reading specialist (clinical subject) together with the simulated case of reading difficulty and attendant data-gathering materials. The experimenter's major tasks were to oversee the entire session, including providing necessary directions, presenting all simulated case materials requested by the subject as well as providing other material required for data collection, audio recording selected portions of the sessions, and asking questions of the clinical subject at specified times during the session. The clinical subject's major tasks were to collect information concerning the case as she normally would, prepare a written

diagnosis and remediation, and explicitly pair remediation with prior diagnoses.

Design of the Study

The design of the study called for random assignment of subjects to SIMCASES. Each simulated case was based on a real child who was served by the Reading Clinic at Michigan State University. The problems represented in the four cases (all were males) were considered to be representative of those commonly encountered in the public schools. For each SIMCASE a cue inventory was provided listing the data (cues) available for that case (Appendix A).

Each SIMCASE had its equivalent form--a superficially disguised replicate of the original prepared by making minor changes in the original data base, and randomly re-ordering the cue inventory (Lee and Weinshank, 1978).

The data in any given SIMCASE included: information about achievement tests, family background, cognitive ability, group and individual reading diagnostic measures, classroom information, etc.

The information was presented in a variety of formats: test booklets, audio tapes, examiner's comments, and test scores.

All cases contained a taped audio interview with the student, a brief statement of the reason for referral to the Reading Clinic, and an artist's sketch of the child based on hearing a taped interview.

The overall design of the study--random assignment of subjects to cases and cue inventory forms--is shown in Table 1.

Table 1: Overall Design of the Study

Teacher	Session 1	Session 2	Session 3
A	Case 1, F.2	Case 4, F.1	Case 1 (Replicate) F.2
B	Case 2, F.2	Case 3, F.1	Case 2 (Replicate) F.2
C	Case 3, F.1	Case 1, F.2	Case 3 (Replicate) F.1
D	Case 4, F.1	Case 2, F.2	Case 4 (Replicate) F.1
E	Case 4, F.1	Case 2, F.2	Case 4 (Replicate) F.1
F	Case 3, F.1	Case 1, F.2	Case 3 (Replicate) F.1
G	Case 2, F.2	Case 3, F.1	Case 2 (Replicate) F.2
H	Case 1, F.2	Case 4, F.1	Case 1 (Replicate) F.2

F.1 and F.2 indicate which of the two forms of cue inventory was used.

Table 2 summarizes the content of each session. A minimum of one week separated sessions one and three.

Table 2: Summary of Contents of Sessions

Session 1	Session 2	Session 3
Work with one of the four original SIMCASES	Work with one of the remaining three original SIMCASES	Work with equivalent form (replicate) of the Session 1 SIMCASE

The Subjects

The eight subjects in the study (all females) were teachers who had an earned master's degree in reading and a minimum of four years of reading-related teaching experience.

In order to account for the possible influence of academic training on diagnostic/remedial performance, two groups of specialists were chosen: those who received their entire undergraduate and graduate training at universities in Michigan and those who received their entire undergraduate and graduate training at universities in the Chicago, Illinois area. There were four reading specialists in each group.

Since data collection was a lengthy procedure requiring three three-hour sessions for each subject, all teachers received payment of \$90.00 for their participation in the study. The 12 sessions with the Chicago-area subjects were conducted over a three week period at the Department of Reading, Northeastern Illinois University, Chicago. The 12 sessions with the greater Lansing area subjects were conducted at the College of Education, Michigan State University, East Lansing.

None of the clinicians was familiar with the previous work conducted by the Clinical Studies Group of the Institute for Research on Teaching. A profile of the teachers who participated in the study is presented in Table 3.

Method

The procedures for sessions one and two were identical. Session three differed only in that a final debriefing was added during which the subjects were asked to reflect on some of their decisions.

The task of the subjects in all sessions was to diagnose the SIMCASE, propose an initial remediation plan, and associate remedial and diagnostic statements.

Table 3: Profile of Participating Subjects

Teacher	Yrs. of Teaching	Most Recent Degree	Grades Taught	Current Responsibilities
1	10	Ed. Specialist	Elem./Jr.High	Reading Consultant
2	18	Masters	Elem./Jr.High	Reading Specialist
3	7	Masters	Elementary	Supportive Reading
4	27	Masters	Elementary	Remedial Reading
5	9	Masters	Elem./Jr.High	Reading/ Lang. Arts
6	4	Masters	High School	Remedial Reading
7	8	Masters	Primary	Reading Resource
8	9	Masters	Primary	All Subjects

The sequence of instructions is detailed in Appendix B. Briefly:

1. The teacher was given an introduction about the study, followed by instructions regarding payment and granting informed consent. During the first session only, a background questionnaire was completed by the subject. An overview of the three sessions was read and the subject, using a training SIMCASE, practiced using the cue inventory to request information about the case.
2. After the brief practice session, the teacher was shown the referral statement and sketch for her assigned case and

listened to the taped interview. She was then allowed up to thirty minutes to request information using the cue inventory, taking notes if she wished. The cue requests were entered on a standard form together with time of request (Appendix C). The session was audio taped as well. When the thirty minutes were over, the teacher was given up to twenty minutes to write her diagnosis, and up to another twenty minutes to write an initial remediation plan. Wide-lined paper was provided, and all notes made and cues collected were at hand while the diagnosis and remediation were being written.

3. After a short break, the subject worked with her written diagnosis according to the following protocol:
 - (a) Using a sample diagnosis as a guide (Appendix D) the subject was given an opportunity to practice circling, numbering, and coding diagnostic statements. The codes were: S for strength, W for weakness, and Obs. for observation. (An observation was defined as a neutral statement which characterized the case, e.g., the child has two siblings.)
 - (b) After completing the practice sample, the subject repeated the task with her own diagnosis.
 - (c) Using a sample diagnostic checklist as a guide, the subject practiced the procedure for transferring the sample numbered, coded diagnostic statements to the checklist.

- (d) After completing the practice, the subject repeated the task with her own diagnosis using a fresh diagnostic checklist (Appendix E).

The protocol for working with the written remediation paralleled that for the diagnosis with two differences:

- (a) A sample remediation was used to practice circling, numbering, and coding remedial statements (Appendix F).

The codes were: S for strength, W for weakness, Obs. for observation, and T for treatment. (b) Numbered, coded remedial statements were transferred to a remedial checklist (Appendix G).

4. After the diagnosis and remediation checklists were completed the clinician directed her attention to her written diagnosis while the experimenter read aloud to the subject each numbered remedial element in turn. For each one, the subject indicated which (if any) diagnostic elements she associated with the remedial element. She then read the number(s) and code(s) of each associated diagnostic element. The experimenter wrote down all associations on a standard form (Appendix H), which was formatted differently for use in Session 3 only (Appendix I). The session was audio taped.

Unmatched remedial element numbers and unmatched diagnostic element numbers were entered on a separate form (Appendix J).

Sessions 1 and 2 ended at this point. Session 3, however, continued from this point, following a short break.

5. To complete session 3 the clinician was asked to reflect on the recommendation/diagnosis associations just completed. Two questions were asked about each association: (1) Why did you choose this remediation?, and (2) What is the association between the remediation and the diagnosis? The questions and alternative responses were listed on a guide sheet (Appendix K), and the experimenter noted responses on the remediation/diagnosis association form used in Session 3.

If any remedial elements were left unassociated the subject was asked to look at the written remediation. The following statement was made by the experimenter for each remediation: Remedial element.....is not associated with a diagnostic element because..... The statement and alternative responses were listed on a guide sheet and the experimenter noted responses on a standard form (Appendix L).

If any diagnostic elements were left unassociated, the subject was asked to look at her written diagnosis. The following statement was made by the experimenter for each diagnosis: Diagnostic element number.....is not associated with a remedial element because..... The statement and alternative responses were listed on a guide sheet and the experimenter noted responses on a standard form (Appendix M).

The entire debriefing was audio taped and the subject was given the opportunity to comment on all responses made.

6. To close the session, the subject was asked to respond to three open-ended questions.

- (a) What guides you in gathering information about a child and in determining what should be written in the diagnosis?
- (b) What do you see as being the relationship between diagnosis and initial treatment plan?
- (c) Do you have any comments about the session?

All responses were audio-taped.

Data Analysis

During the course of each of the 24 observational sessions, subjects collected case information (cues), wrote a diagnosis and remediation and transferred both to standardized checklists, and associated remedial with diagnostic statements. These procedures resulted in five products being collected for each session: a cue list, diagnosis, remediation, remediated diagnoses, and specific remedial/diagnostic associations. These products were analyzed on a case-by-case basis (six sessions for each case). The data was to be analyzed in order to determine:

1. Agreement among and between clinicians on
 - (a) Diagnostic statements
 - (b) Remedial statements
 - (c) Information collected
 - (d) Remediated diagnostic statements, and
 - (e) Specific remedial/diagnostic associations;
2. Strength of correlations between various products; and
3. Any differences in performance between clinicians trained in different institutions.

The standardized checklists were the source of data for the analyses. They provided a common metric for making comparisons across sessions. The diagnostic and remedial checklists were derived from diagnostic and remedial statements made by reading specialists during the course of previous observational studies conducted by the Clinical Studies Group. These statements were categorized and redundancies eliminated. There were 101 statements in both lists grouped into a number of categories. Each diagnostic statement could be classified as a strength, weakness, or observation. Each remedial statement could be classified as a strength, weakness, observation, or treatment.

Statements from the written diagnosis and remediation which subjects felt could not be matched to domain statements were written in, either under "Other" within one of the existing categories, if possible, or under a separate major category labelled "OTHER." Subjects were asked to make as limited use as possible of "Other/OTHER" statements. The formal analyses did not include these statements since they were not a part of the fixed diagnosis/remediation domains. Analysis of "Other/OTHER" statements was dealt with separately.

The cue domain for each case was simply a sequentially numbered list of all the cues present in that case. Each cue requested during a session was noted. In preparing data sheets, the appropriate domain number was placed next to the requested cue.

The agreement statistics were calculated using (1) a computer statistical analysis system (Observational Study Data Analysis System)

developed and maintained by the Institute for Research on Teaching, Michigan State University, and (2) an interactive statistical analysis system (MIDAS) at the University of Michigan.

A random sampling of data was computed by hand across numerous observational studies, and the system was found to be operating accurately.

Agreement Statistics

(1) Proportional Agreement

In order to determine the extent of agreement for each diagnosis/remediation/cue for a given case, a proportion was computed for each statement in the domain for diagnosis, remediation, and cues. If a statement received a value of zero (0.00), it was not mentioned during any of the six sessions for the case. If a statement received a value of one (1.00), it was mentioned during all six sessions.

For example, if the statement, "Basic Sight Words, Weakness," was mentioned in three of the six sessions, the proportional agreement would be:

$$P.A. = \frac{\text{No. of sessions during which the statement was made}}{\text{Total no. of sessions}}$$

$$P.A. = \frac{3}{6}$$

$$P.A. = 0.50$$

(2) Commonality

The clinician's products in a session were compared with those from the other five sessions devoted to the case in order to assess how similar was one clinician's performance to the total group.

Any statement made in two or more sessions (out of six possible) was included in the domain used for calculating commonality. For example, assume there were twenty statements in a diagnostic domain for a given case. A clinician who had mentioned fifteen of those statements in her individual diagnosis would have a higher commonality score than one who had mentioned only six of those twenty domain statements. The performance of the first clinician would be more similar to that of the group than that of the second clinician.

(3) Intercorrelations

a) Inter-Clinician Correlation. The diagnostic/remedial/cue (Dx/Rx/Cx) categories mentioned by one clinician were compared with those mentioned by a second clinician for the same case. A Phi correlation, a measure of inter-clinician agreement, was computed for each pair of clinicians (Appendix N). The comparison is summarized in the grid below.

b) Intra-Clinician Correlation. The diagnostic/remedial/cue categories mentioned by the clinician at time one (T1) were compared with those mentioned by the same clinician for the replicate form of the same case at time two (T2). A Phi correlation, a measure of intra-clinician agreement, was computed for each pair of sessions for given

clinicians (Appendix N). The comparison is summarized in the grid below.

Clinician A, SIMCASE Q			
		PRESENT (+)	ABSENT (-)
Clinician B, SIMCASE Q	PRESENT (+)	Frequency count of statements in the domain present in both clinicians' sessions Diagnosis/Remediation/Cues A	Frequency count of statements in the domain present in clinician B's session but not in clinician A's Diagnosis/Remediation/Cues B
	ABSENT (-)	Frequency count of statements in the domain present in clinician A's session but not in B's Diagnosis/Remediation/Cues C	Frequency count of statements in the domain absent in both clinicians' Diagnosis/Remediation/Cues D

The presence of a large percentage of statements (more than 85%) in the "D" cell (the statement is absent in both sessions) artificially inflated the intercorrelations since it represented, in effect, agreeing to disagree. A statistic developed by Professor A. Porter (Institute for Research on Teaching, Michigan State University) was designed to correct for this occurrence, by including in the computation only the values in the A, B, and C cells $\frac{A}{A + B + C}$. All intercorrelations were calculated twice (including and excluding the D cell) so that the impact of cell size could be seen.

Clinician i SIMCASE Q, Form One

PRESENT (+)

ABSENT (-)

Clinician i SIMCASE Q,

Form Two

Frequency count of statements present in the domain in both sessions for form one and form two of SIMCASE A	Frequency count of statements in the domain present in SIMCASE form two but not in SIMCASE form one B
Frequency count of statements in the domain present in the session for SIMCASE form one but not SIMCASE form two C	Absent in both sessions for form one and form two of SIMCASE D

Additional Analytic Procedures

Pearson Product Moment Correlations were computed for various pairs of products in order to determine whether performance on one product could be considered to be predictive of performance on another.

Finally, t-tests for differences between means were computed to establish if the two groups of subjects, Illinois and Michigan, performed significantly differently.

Summary

The primary objective of this study was to establish the nature and extent of the relationship between diagnosis and remediation in reading. To obtain the necessary data, eight experienced, credentialed, practicing reading specialists, four from Illinois and four from Michigan, each participated in three sessions during which they

performed certain tasks using simulated cases of reading difficulty. Two of the cases were equivalent forms of the same problem. The third represented a different problem.

The products that resulted from each session were analyzed to determine (1) extent of agreement among and between clinicians (proportional agreement, commonality, intercorrelations); (2) strength of correlations between products; and (3) any differences in performance between Illinois-trained and Michigan-trained subjects.

CHAPTER FOUR

RESULTS

Introduction

The major purpose of this study was to provide an objective measure of the nature and extent of the association between diagnosis and remediation in reading.

Eight reading specialists working in a controlled setting provided the data to be described in this chapter. Specifically, in a total of twenty-four sessions they requested information from a simulated case of reading difficulty and:

1. wrote a diagnosis for a SIMCASE of reading difficulty;
2. wrote an initial remediation plan;
3. transferred diagnostic statements to a standardized diagnostic checklist and remedial statements to a standardized remedial checklist;
4. made associations between remedial and diagnostic statements;
5. reported on reasons for making the associations;
6. reported on reasons for the presence of unassociated diagnostic elements (if any) and unassociated remedial elements (if any);
and
7. discussed their conduct of the diagnosis/remediation procedure.

The data analytic results will be reported first for the three kinds of agreement described in Chapter Three:

1. Proportional agreement (the number of times a statement was mentioned across the six sessions for each case);
2. Commonality (the extent to which an individual mentions the same statements as did the group for a given case); and
3. Intercorrelations (the extent of agreement on statements characterizing a given case between a clinician with herself and with another clinician).

Further results to be reported are Pearson Product Moment Correlations between products (e.g., between diagnosis and remediation, cues and diagnoses, etc.) and t-tests for differences in mean performance between Illinois and Michigan subjects.

The reading specialists in this study examined four simulated cases of reading difficulty which presented information about real children with reading difficulties. The children's grade placement ranged from grades three to seven.

Case One was a third grader with (1) a significantly depressed sight vocabulary, (2) a serious problem with decoded word recognition skills, (3) inadequate fluent message segmentation, and (4) no significant problem comprehending anything he could read or listen to.

Case Two presented a sixth grader who exhibited (1) significantly retarded sight vocabulary and word recognition behaviors, (2) a significant problem with analysis on multisyllabic words coupled with poor visual segmentation and high frequency hearing loss, (3) difficulty phrasing material of higher level semantic content, and

(4) comprehension problems related to the demands of higher level content-related materials.

Case Three presented a seventh grader with (1) reasonably intact sight vocabulary and word recognition performance, (2) significant problems with higher level decoding skills and their application during reading, (3) a serious problem with fluent texting of material, and (4) a comprehension problem specifically related to the meaning demands of content-related materials.

Case Four was a fourth grader who showed (1) a significantly depressed sight vocabulary, (2) a significant problem with the learning and application of decoded word recognition skills, (3) no major problem with text segmenting, and (4) no language-based comprehension problem.

In order to give the reader a sense of the data used for comparing clinician performance, several diagnoses and remediations written by clinical subjects in the study are presented for Case Three. Clinician A's reliability (her agreement with herself on the same case) was in the lowest third of the range; clinician B's was in the highest. For convenience, their numbered diagnostic and remedial statements are presented serially instead of in the prose form in which they were actually written. The code letters assigned to each diagnostic statement are: (s) for strength, (w) for weakness, (obs) for observations, and (t) for treatment.

A master diagnosis and remediation plan compiled by a senior reading clinician on the faculty of Michigan State University is presented in Appendix O.

Diagnosis: Clinician A

1. He has poor study skills. (w)
2. His lack of concentration and interest could have led to his poor study skills. (w)
3. He is not well liked by students or adults. (w)
4. He lacks motor control. (w)
5. He has a history of speech problems. (w)
6. His oral reading was performed word by word. (w)
7. He repeated phrases. (w)
8. He also had irregular pauses. (w)
9. His tone was a sing-song type and sometimes choppy. (w)
10. This made a very unpleasant auditory interpretation of the selection. (obs)
11. His attitude towards reading aloud was negative. (w)
12. His comprehension during his oral performance was good. (s)
13. He did not encounter much difficulty until selection 8. (s)
14. His I.Q. was average. (s)
15. His silent reading was lacking for his score was a high fourth grade. (w)
16. His listening comprehension was at grade level. (s)
17. His word recognition was almost at grade level. (s)
18. His word analysis was almost at grade level. (s)
19. His spelling was a low third grade level. (w)
20. His spelling problems were in all three areas: syllabication, (w)
21. vowels, (w)
22. and consonant substitutions and omissions. (w)

23. His auditory test indicated some hearing loss. (w)
24. This may have led to some of his speech and spelling problems.
(obs)
25. Even though he has poor motor skills, his main interest is sports.
(obs)
26. His poor motor control could be the cause of his poor and
irregular handwriting. (w)
27. His word analysis was nearly at grade level. (s)
28. This indicated he was better at auditorily producing the sounds
rather than being able to write the letter names for the sounds.
(obs)
29. He used beginning sound clues when encountering unknown words,
(s)
30. but often missed medial and ending syllables. (w)

Remediation: Clinician A

1. First we are going to have to work on his attitude towards
reading. (t)
2. Providing him with purpose and worth for reading will help. (t)
3. Because he had a desire to lead and win friendship, maybe some
instruction on reading for information so he could converse and be
interesting to others will help. (t)
4. Reading topics of his interest and others who are close to him will
also help in providing a purpose for reading. (t)
5. Because of the unpleasant auditory interpretation of selections,
he should also work on oral reading skills. (w)

6. Reading selections silently and thinking about the subject matter they contain may help. (t)
7. Listening to tapes and then following the listening session trying to imitate their expressions and phrasing would also help. (t)
8. Practicing a selection so as to be able to perform it as a dramatic reading may also help prevent his uninterested tone while reading aloud. (t)
9. More practice with silent reading would help. (t)
10. Questions preceding the silent reading will motivate him to be reading for information. (t)
11. Spelling was a severe deficiency. (w)
12. Cards containing the correct letters in misspelled words and spaces where the incorrect letters were would get him to think about the unknown parts of words rather than studying the whole word. (t)

Diagnosis: Clinician B

1. His is of average I.Q. (s)
2. He comes from a good home. (s)
3. He appears to have a normal sibling relationship. (s)
4. He suffers from both an auditory (w)
5. and visual problem. (w)
6. His auditory test showed a marked acuity problem in the left ear. (w)
7. His visual test showed that he was farsighted. (w)
8. Emotionally he seems to suffer from a school problem. (w)

9. His teacher felt that no one liked him, and she had very negative comments about him. (w)
10. He was aggressive and was generally found to be not too pleasant a person. (w)
11. The parents didn't see a social problem for him. (obs)
12. They were more concerned with a speech (obs)
13. and a motor-coordination problem. (obs)
14. They claimed he was shy which doesn't quite agree with his teacher's report. (obs)
15. He tested on the Durrell at mid high fourth grade on the oral (w)
16. and a high fourth on the silent portion. (w)
17. His comprehension oral was good. (s)
18. His written comprehension was good, and the examiner noted good recall. (s)
19. He frustrated at the sixth grade level. (w)
20. His rate lagged behind on both the Durrell and the Gates. (w)
21. His word recognition was excellent. (s)
22. He had some problems on base words. (w)
23. Otherwise, his recognition of blends (s)
24. and syllables was good. (s)
25. His main problem seems to be one of comprehension (w)
26. and rate. (w)
27. He could be a non-context reader as evidenced by the high vocabulary score on the Gates. (w)
28. He is not phrasing well. (w)
29. This could mean he is not reading for context. (w)

Remediation: Clinician B

1. In remediating this problem first get a complete visual (t)
2. and hearing test readministered. (t)
3. Check into his school situation. (t)
4. and determine how he is doing socially with his peers. (obs)
5. School relationships are very important when we look at academic progress. (obs)
6. Also check the home situation and discuss the school problem with the parents. (t)
7. Remediation should center on comprehension. (t)
8. He is approximately two years below level (w)
9. but is not too deficient in word recognition skills. (s)
10. Since his recall was good at fourth grade level, (s)
11. I would work on higher level comprehension skills. (t)
12. Give him books of interest and at a readability he can cope with. (t)
13. Do not ignore word recognition but do not focus on it. (t)
14. He had some problems in structural analysis (w)
15. and in recognizing base words, (w)
16. so I would give him work with word parts. (t)
17. Blends could be worked on using various phonics exercises. (t)
18. I would teach him to use context. (t)
19. Give him exercises on reading for a purpose. (t)
20. His rate should improve as he begins reading for a purpose. (obs)
21. Rate must only improve with comprehension. (obs)

22. Various timed exercises with comprehension checks can be useful here. (t)

Proportional Agreement

(1) Diagnosis

Across the four cases, a total of 304 diagnostic statements was made. Analysis of these statements revealed that barely one-third were mentioned in two or more sessions, as shown in Table 4.

Table 4: Proportion of Diagnostic Statements
Mentioned in Two or More Sessions

Case One	0.32
Case Two	0.42
Case Three	0.35
Case Four	0.33

The two sets of categories--those mentioned in two or more sessions and those mentioned only once had different characteristics. The proportion of statements coded as being a weakness (as opposed to strength, or observation) was substantially higher for the categories mentioned in two or more sessions than for those mentioned only once.

Thus, the more idiosyncratic statements which made up the bulk of the diagnostic categories for all the cases tended to revolve more around statements of strengths and observations. Those statements made more often across sessions tended to focus more sharply on problematic aspects of the child's reading performance.

Table 5: Proportion of Weakness Statements
in Set of Diagnostic Statements

	Mentioned Once	Mentioned in Two or More Sessions
Case One	0.32	0.66
Case Two	0.24	0.44
Case Three	0.49	0.69
Case Four	0.34	0.44

Looking across the four cases, 33 specific diagnostic statements were mentioned in two or more sessions. They clustered within nine categories as follows:

1. INTELLECTUAL POTENTIAL (verbal, nonverbal);
2. WORD RECOGNITION (general, basic sight words, use of letter order);
3. ORAL READING (general, rate, fluency, phrasing, punctuation, substitutions);
4. WORD ANALYSIS (general, use of letter/sound associations, syllables, root words, suffixes, ability to blend, use of blends, initial consonants);
5. HEARING (general, acuity, discrimination);
6. VISION (acuity);
7. SILENT READING (rate, comprehension);
8. COMPREHENSION (general, silent, oral, listening); and

9. AFFECTIVE (general, attitude, relationship with peers, emotional adjustment, home environment).

To summarize, about one-third of the diagnostic statements made were agreed upon as being characteristic of a given case. Those statements were more apt to be coded as being weaknesses rather than strengths or observations. Diagnostic statements mentioned only once were drawn from across the entire domain. They were more apt to be coded as strengths or observations (Appendix P). Across all four cases, combinations of 33 diagnostic statements were mentioned more than once (Table 6).

(2) Remediation

A total of 270 remedial statements was made across the four cases. Analysis of these statements showed that the number of remedial statements mentioned in two or more sessions as compared to the total number mentioned for each case was lower than that obtained for diagnostic categories.

Table 7: Proportion of Remedial Statements
Mentioned in Two or More Sessions

Case One	0.34
Case Two	0.40
Case Three	0.15
Case Four	0.12

Table 6: Agreed-Upon Diagnostic Statements Across Cases

Cases				Statements
1	2	3	4	
X	X	X	X	Intellectual potential
X			X	Word recognition: general
X			X	sight
X				reversals
X	X			Oral reading: general
X			X	rate
X				punctuation
X		X		substitutions contextually acceptable
X		X		phrasing
X				fluency
X	X		X	Word analysis: general
	X			suffixes
	X		X	letter/sound associations
			X	initial consonants
X		X	X	blending ability
	X	X		syllables
		X		use of blends
		X		use of root words
	X			Hearing: discrimination
	X			general
	X	X		acuity
		X	X	Vision: acuity
X	X	X	X	Silent reading: rate
	X	X		comprehension
		X	X	Comprehension: oral
		X	X	general
X	X	X	X	silent
		X	X	listening
X				Affective: general
		X	X	attitude toward reading
			X	home environment
		X		relationship with peers
		X		adjustment

Again, as with the diagnoses, the two sets of categories--those mentioned in two or more sessions and those mentioned only once had different characteristics. There were more treatment statements in the set of remedial statements made in two or more sessions than in the set of remedial statements made only once.

Table 8: Proportion of Treatment Statements
in Set of Remedial Statements

	Mentioned Once	Mentioned in Two or More Sessions
Case One	0.34	0.63
Case Two	0.40	0.63
Case Three	0.49	1.00
Case Four	0.43	0.63

The idiosyncratic statements contained a large number of statements which were essentially diagnostic rather than remedial. That is, they were seen as strengths, weaknesses, or observations rather than as treatments. However, as discussed in Chapter Three, the remedial checklist reflects the way in which clinicians with similar training and experience wrote their remediations in previous studies. Those statements made in two or more sessions focused almost exclusively on treatment actions that needed to be taken.

Across the four cases, 27 specific remedial statements were mentioned in two or more sessions. They clustered within seven categories:

1. SIGHT WORDS (identify high frequency sight words in context/isolation; build sight vocabulary; use in language experience stories; use sight words, phrases);
2. STRUCTURAL ANALYSIS (syllables general; count number of syllables seen and heard; add prefixes/suffixes to root words; identify root words);
3. PHONETIC ANALYSIS (general; drill on short/long vowel generalizations; blend sound units into words; apply phonics to attack unknown words);
4. ORAL READING (improve rate/phrasing);
5. VISUAL (memory general; discriminate visually similar words);
6. COMPREHENSION (general; information gathering general; use context);
7. MOTIVATION (provide high interest materials; involve student in setting own purpose).

To summarize, a small number of remedial statements was agreed upon as being appropriate to a given case. These statements were almost always coded as being treatments rather than strengths, weaknesses, or observations. In contrast, those remedial statements mentioned only once were drawn from across the entire domain and were more likely to be coded as strengths, weaknesses, or observations, rather than as treatments (Appendix P). Across all four cases, combinations of 27 remedial statements were mentioned more than once (Table 9).

Impact of additional "Other/OTHER Statements" in the remedial and diagnostic checklists. Both the diagnostic and remedial checklists

Table 9: Agreed-Upon Remedial Statements Across Cases

Cases				Statements
1	2	3	4	
X			X	Sight words: identify high frequency (context)
X	X		X	build sight vocabulary
X				general
X			X	use in language experience
			X	use sight words/phrases
X				Oral reading: improve rate
X				general
X		X		improve fluency
	X		X	Phonetic analysis: general
X				apply phonics to attack new words
X			X	drill on short/long vowel generalizations
			X	blending, general
			X	blend sound units
X				Structural analysis: general
X			X	recognize common word parts
	X			add prefixes/suffixes
	X			count syllables
	X			identify root words
	X			teach syllables
			X	Vision: discriminate visually similar words
	X	X	X	Comprehension: general
	X	X		information gathering
		X		demonstrate factual recall
X	X	X		use context as aid in comprehension
			X	Motivation: general
		X		setting own purpose
	X		X	provide high interest materials

allowed for the addition of "Other/OTHER" statements. These statements were not part of the standardized domains for diagnosis and remediation and were not included in any of the product analyses reported.

The proportion of these statements to those that fit the checklist are shown in Table 10.

Table 10: Proportion of "Other/OTHER" Statements
in Set of Checklist Statements

	Statements	
	Diagnosis	Remediation
Case 1	0.04	0.11
Case 2	0.01	0.07
Case 3	0.09	0.25
Case 4	0.06	0.15

The additional statements in the remedial checklists were examined first in order to determine if they differed in any substantive way from the standardized statements. They did not, as the following analysis indicates.

1. The proportion of treatment statements to statements of strength, weakness, and observation was the same for both the additional and standardized statements.
2. The additional statements were distributed across the same categories as the standardized statements; and

3. All the additional statements were mentioned only once, reflecting the largely idiosyncratic nature of the remediations in general.

The additional statements in the diagnostic checklist also did not differ in any substantive way from the standardized statements.

1. The proportion of weakness statements to statements of strength and observation was the same for the additional and standardized statements;
2. The additional statements were distributed across the same categories as the standardized statements; and
3. The additional statements were all idiosyncratic.

It should be noted also that the existing major categories in both checklists accommodated almost all of the additional statements. Only a very small number fell wholly outside the framework of either checklist's categories. "OTHER" statements made for the diagnostic checklist were (1) spelling, (2) handwriting, (3) writing vowel sounds, (4) writing consonant sounds, and (5) writing syllables. All were coded as a weakness. "OTHER" statements made for the remedial checklist were (1) vocabulary meaning in context, (2) write daily log, (3) silent reading, (4) emotional, (5) finger pointing, (6) lip reading, (7) peer relationships, and (8) parent educational background. All but the last two statements were coded as treatment.

(3) Cues

The domains for diagnosis and remediation are all-inclusive, i.e., each domain serves all four cases. In contrast, each case has its own cue domain with the size of the domains ranging from 83 to 104 items.

The number of cues requested in two or more sessions as compared to the total number requested for each case was relatively high.

Table 11: Proportion of Cues Mentioned in Two or More Sessions

Case 1	0.79
Case 2	0.71
Case 3	0.63
Case 4	0.68

As with the diagnoses and remediations, the two sets of cue categories--those requested in two or more sessions and those requested only once--had different characteristics. The least agreed upon cue forms (mentioned only once) were those that presented information in the relatively high inference form of test booklets and audio tapes which required that the subjects make judgments about the child's problems based on an examination of the child's performance on the protocols.

Table 12: Proportion of Requests for Test Booklets/Audio Tapes

	Requested Once	Requested in Two or More Sessions
Case 1	0.54	0.14
Case 2	0.50	0.27
Case 3	0.31	0.11
Case 4	0.37	0.20

The most agreed-upon cue forms were those that provided information in the relatively low inference form of test scores and comments (Appendix P). The subjects simply had to note the score or read the comment.

The test scores and comments requested in two or more sessions dealt with:

1. Background (home/school);
2. Cognitive ability (Weschler, Slosson, Peabody);
3. Diagnostic tests of reading ability including oral reading, word analysis, silent reading, visual memory, etc. (Durrell, Gates, McKillop, Individual Reading Analysis);
4. Graded word lists (Slosson Oral Reading Test);
5. Basic sight vocabulary (Dolch);
6. Reading achievement tests (Gates, McGinitie);
7. Visual/auditory reports; and
8. Auditory discrimination (Wepman).

Some clinicians collected many items of information within each category, others collected very few items. Thoroughness of collection had minimal effect on diagnostic reliability, however. The clinician who collected the fewest items of information ($\bar{X} = 15$) had zero diagnostic reliability. The clinician who collected the most items ($\bar{X} = 44$) had a diagnostic reliability of only 0.12.

The categories of cues were collected consistently across cases. However, agreement within cases on specific diagnostic and remedial statements which characterized that case was extremely limited. For a given case, then, the interpretation of commonly selected cues

appeared to be highly idiosyncratic, with little agreement as to what to do with the data once it has been collected.

In summary, the most frequently requested cues were in the low inference form (i.e., the child's work itself was not available for inspection) of test scores and comments dealing with word recognition and analysis, cognitive ability, oral and silent reading, hearing, vision, and background. The most frequently mentioned diagnoses and remediations mirrored these cues, with the interesting exception of background information which did not seem to have any impact across sessions. It appears, then, that across sessions, common cues resulted in largely non-common diagnoses (primarily problems) and non-common remediations (primarily treatments), indicating a lack of mutually agreed upon interpretations of common information sources.

(4) Remediated Diagnoses

Not all diagnostic statements were remediated, as Table 13 illustrates.

On the average, roughly two-thirds of the diagnostic statements made across cases were remediated. Diagnostic statements coded as weaknesses were more likely to be given remediation than those coded as either strengths or observations. This held both for remediated diagnoses mentioned once as well as for those mentioned in two or more sessions (Appendix P).

As with the diagnostic and remedial statements, those diagnoses seen as requiring treatment were usually not agreed upon across two or more sessions.

Table 13: Proportion of Remediated Diagnoses
to Total Diagnostic Statements

	Total Number of Diagnostic Statements Made	Statements Number of Diagnoses that Were Linked to a Remediation	Proportion of Diagnostic Statements that Were Remediated
Case 1	74	45	0.60
Case 2	64	45	0.70
Case 3	91	58	0.60
Case 4	75	51	0.66

It should be noted that, for all cases, it was a subset of the more agreed-upon diagnoses that were chosen for remediation. That is, the clinicians did not remediate idiosyncratic diagnostic statements. Rather, they remediated only those diagnostic statements that analysis showed to have been repeated in two or more sessions and which were shown to be grouped under nine categories.

Table 14: Proportion of Remediated Diagnoses
Mentioned in Two or More Sessions

Case 1	0.29
Case 2	0.37
Case 3	0.17
Case 4	0.41

(5) Specific Remedial/Diagnostic Associations

Across the four cases, only 31 specific remedial/diagnostic associations (out of a total of 654 made) were repeated across two or more sessions. The proportion of agreed-upon associations to all associations made is shown in Table 15.

Table 15: Proportion of Agreed Upon Remedial/Diagnostic Associations (Two or More Sessions)

Case 1	0.02
Case 2	0.06
Case 3	0.02
Case 4	0.05

Put another way, from 94 percent to 98 percent of all specific remedial-diagnostic associations were idiosyncratic to a given session, that is, they were mentioned only once. The few associations that were made more than once were more likely to pair treatment with weakness than were the idiosyncratic remedial/diagnostic associations (Appendix P). Table 16 shows the proportion of treatment/weakness associations for single sessions and for two or more sessions.

An average of 17 percent of the associations made during a given session paired a single remediation with more than one diagnostic statement. A treatment would be paired with different weaknesses or with strengths and observations as well as weaknesses. For example, one remedial statement coded T (the student requires instruction in a systematic phonics program) was associated with three diagnostic

Table 16: Proportion of Treatment/Weakness Associations
in Set of All Remedial/Diagnostic Associations

	Proportion of Treatment/ Weakness Associations Mentioned During One Session Only	Proportion of Treatment/ Weakness Associations Mentioned During Two or More Sessions
Case 1	0.40	1.00
Case 2	0.31	0.50
Case 3	0.43	1.00
Case 4	0.41	0.66

statements coded, respectively, W, S, and Obs. (The student is practically a non-speller. He can hear sounds in words. He needs instruction in word analysis.)

The effect of this multiple pairing was to generate a large number of associations for each session. As was shown, virtually none of them was ever repeated in another session.

The results of the analyses for proportional agreement can now be summarized.

1. A small proportion of the total diagnostic and remedial statements mentioned for a case were repeated across two or more sessions. Most statements were idiosyncratic, that is, they were mentioned in only one session.
2. Diagnostic statements coded as "Weakness" and remedial statements coded as "Treatment" were more likely to be agreed upon across sessions than were diagnostic statements of strength

and observation and remedial statements of strength, weakness, and observation.

3. Roughly two-thirds of all diagnostic statements were remediated. However, of these, only a small percentage were agreed upon across two or more sessions. Diagnoses that were remediated were almost always coded as a weakness.
4. There was considerably more agreement on which cues should be collected than on the diagnoses and remediations which resulted from examination of these cues.
5. Associations between remediation and diagnosis mentioned two or more times, though very few in number, were more likely to pair treatment with weakness than were remediation/diagnosis associations mentioned only once.

The data partially supported two hypotheses:

1. Statements of weakness were repeated more reliably across sessions than were all diagnostic categories; and
2. Statements of weakness were more often associated with treatments across sessions.

The hypothesis dealing with remediated diagnostic statements--that they are repeated more reliably across sessions than are all diagnostic statements--was not supported by the data. While remediated diagnoses were more likely to be categorized as a weakness, agreement as to specifically which ones would be remediated was lower than agreement for all diagnostic categories.

As the following summary figure and tables show, most statements were made only once. Even setting the most generous standard for

reliability (statement is mentioned in two of the six sessions) resulted in agreement only on one-third or fewer of the total statements made. When the standard for reliability is changed to reflect those statements made in three or more of the sessions, agreement plunges precipitously (Tables 18, 19).

Idiosyncratic statements, therefore, resulted in products that were much longer than was necessary for reliable judgments.

A number of possible sources for the length of individual products was hypothesized: (1) diagnostic statements of strength and observation tended to be more idiosyncratic than were statements of weakness; (2) remedial statements of strength, weakness and observation tended to be more idiosyncratic than were statements of treatment; (3) redundant associations, and associations which paired diagnostic statements other than "weakness" with remedial statements other than "treatment" combined to cause an inflated number of associations; (4) highly idiosyncratic and inconsistent interpretations of background information, test booklets, and audio tapes acted as distractors to the main task of isolating weaknesses and specifying treatments. Use of these materials was shown to be strongly related to (1) both increased use of strength and observation statements within sessions and (2) decreased agreement of statements across sessions.

The data on proportional agreement are summarized in Figure 1 and Tables 17, 18, and 19.

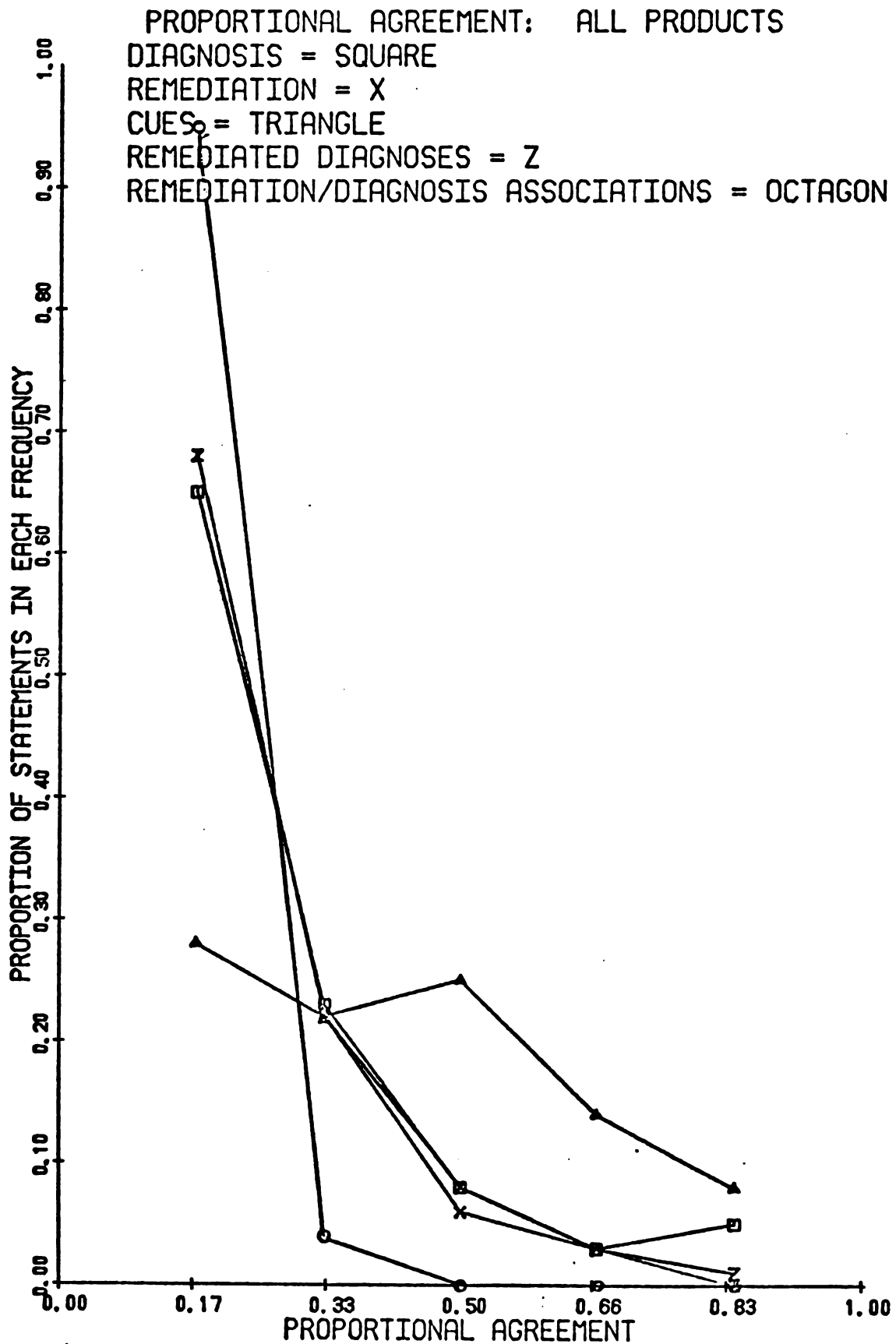


Table 17: Proportions of Idiosyncratic and Agreed-Upon Statements Across Products

	Diagnosis	Remediation	Cues	Remediated Diagnosis	Remedial/Diagnostic Associations
Proportion of total statements made during one session	0.65	0.75	0.30	0.69	0.95
Proportion of total statements repeated over two or more sessions	0.35	0.25	0.70	0.31	0.05

Table 18: Summary Table
Proportion of Statements for Each Session Across Products

	1 Session (0.17)	2 Sessions (0.33)	3 Sessions (0.50)	4 Sessions (0.67)	5 Sessions (0.84)	6 Sessions (1.00)
Diagnosis	0.65	0.23	0.08	0.03	0.01	0.003
Remediation	0.69	0.21	0.06	0.03	0.003	0.003
Cues	0.29	0.21	0.25	0.14	0.08	0.02
Remediated Diagnosis	0.69	0.21	0.07	0.03	0.01	0.00
Remedial Diagnostic Associations	0.95	0.04	0.004	0.001	0.00	0.00

Table 19: Total Number of Statements for Each Session Across Products

	1 Session (0.17)	2 Sessions (0.33)	3 Sessions (0.50)	4 Sessions (0.67)	5 Sessions (0.84)	6 Sessions (1.00)	Total
Diagnosis	196	69	25	9	4	1	304
Remediation	187	58	16	7	1	1	270
Cues	76	55	65	37	20	5	258
Remediated Diagnosis	138	41	13	5	2	0	199
Remedial Diagnostic Associations	623	27	3	1	0	0	654

Characteristics of Unassociated
Remedial and Diagnostic Statements

The results of the pilot study (Chapter Two) indicated that clinicians did not always make associations between diagnostic and remedial statements. The inconsistency was hypothesized to arise from two sources:

1. If a problem in the diagnosis was not associated with any remediation, some possible explanations were that the clinician forgot to include a treatment for the problem and/or the clinician did not know of any appropriate treatment and/or treatment was not called for.
2. If a treatment in the remediation was not associated with any problem in the diagnosis, some possible explanations were that the clinician never articulated the problem in the diagnosis and/or always used this treatment irrespective of diagnosis.

The clinicians in this study had unassociated diagnostic and remedial statements. Across all cases, unmatched diagnostic statements were much more likely to be coded as strengths and observations rather than as weaknesses. This complemented earlier findings that it was statements of weaknesses that were repeated more frequently across sessions, were remediated more frequently, and were associated more frequently with treatments than were either statements of strengths or observations.

During the debriefing at the end of session three, the subjects were asked about the diagnostic statements which were not associated .

Table 20: Unmatched Diagnostic Statements

	Number of Unmatched Strength/Observation Statements	Number of Unmatched Weakness Statements
Case 1	39	13
Case 2	32	8
Case 3	40	37
Case 4	35	6

with a remediation. They were given a guide sheet to help them reflect on the question as follows:

This diagnostic element is not associated with a remediation because:

- a. The diagnostic element cannot be remediated;
- b. The diagnostic element does not require remediations;
- c. I don't know a reason for this diagnostic element;
- d. I forgot to include the remediation in my write up;
- e. other.

Across all the cases, strengths were not associated with remediations almost entirely because of reasons stated in statements a and b above: the diagnostic statement couldn't be remediated or didn't need one. Only twice was an unassociated strength attributed to d above, that is, the clinician had intended to associate it with a remediation but forgot to write the remediation in the initial plan.

Statements of observations also were not remediated because of reasons stated in statements b and d above. Only once was an

unassociated observation attributed to c above: that is, a remediation was seen as necessary but the subject did not know of one.

Weaknesses were not associated with remediations almost exclusively because of reasons stated in c and d above: the subject knew no remediation or forgot to include the remediation in the write-up. In two cases the subject felt that the diagnostic element could not be remediated (a above).

In summary, some diagnostic statements of strength, weakness, and observation in this study were always left unassociated with remediation. Statements of strength and observation were unassociated primarily because they could not be remediated or did not require remediation. Unassociated statements of weakness were left unmatched primarily because the subject meant to provide a remediation but forgot to write it down, and secondarily because the subject simply did not know a remediation for the diagnostic element.

Turning to the characteristics of unassociated remedial statements, the data showed that treatment statements made up by far the major portion of all remediations. Not surprisingly, then, almost all unassociated remediations were statements of treatment. Additionally, the design of the study was such that diagnoses had to be matched to each remediation in turn, making it less likely that any given remediation would go unnoticed.

Total Unmatched Remedial Statements (Out of 260 Statements)

Treatment	= 26
Strength	= 2
Weakness	= 3
Observation	= 1

During the debriefing at the end of session three, the subjects were asked to consider why these remedial statements were not associated with a diagnostic statement. They were given a guide sheet as follows:

- a. This remediation should always be included no matter what the specific diagnosis may be;
- b. The remediation is always appropriate for this type of student;
- c. I forgot to include the diagnostic element in the write-up;
- d. Other.

Treatments were not associated with diagnoses for only two reported reasons: (a) the remediation should always be included and (b) it is always appropriate. In these relatively few instances of unassociated treatments, then, a diagnosis would actually be unnecessary. Irrespective of diagnostic findings, these treatments were seen as always being necessary and appropriate for children with reading difficulty.

The two sources hypothesized to account for inconsistency in associating initial remediation plan with diagnosis (Chapter One) were supported by the data.

1. Statements in the diagnosis were not associated with any treatment because
 - a. they could not be remediated,
 - b. they did not require remediation,
 - c. the clinician meant to provide a remediation but forgot to write it down, and

- d. the clinician knew no remediation for the diagnostic statement.
2. Statements in the remediation were not associated with any diagnostic statements because the clinician believed that the remediation should always be included.

Characteristics of Associated
Remedial and Diagnostic Statements

Most remedial statements were associated with diagnostic statements. During the debriefing at the end of session three, the subjects were asked (1) why they chose the particular remediation they had written down, and (2) what they saw as the association between that remediation and the diagnoses which they associated with it. They used a guide sheet as follows:

1. Why did you choose this remediation?
 - a. I am most familiar with it
 - b. It is usually available
 - c. I usually have success with it
 - d. Other
2. What is the association between remediation and the diagnosis?
 - a. The remediation is effective for this weakness
 - b. The remediation is effective for this strength
 - c. The remediation is effective for observations of nature
 - d. Other

In answering question one ("Why did you choose this remediation?") the subjects indicated that for those remediations coded as treatments, they chose the remediation in question primarily because it was usually

available and usually effective, and secondarily because of familiarity. In other words, the teachers may have been familiar with a wide range of remedial materials and techniques, but they reported usage on the basis of availability and successful prior use.

For non-treatment remediations (strength, weakness, or observation) the subjects had to use the "other" category to explain their choice. Typical comments concerning a non-treatment remediation were, "It's not a remediation, it's an observation (or strength, or weakness)." Some variations were, "I'm not suggesting a treatment, I'm commenting on a weakness;" or, "I'm just noting scores (or comments, or background data, etc.)."

It appeared from these limited debriefing data that the subjects saw a remedial plan as having two categories of statements: remediations (i.e., treatments) and other statements whose purpose seemed to be to reiterate statements made in the diagnosis. These non-treatment statements, as had been demonstrated earlier, were almost always idiosyncratic in nature and served to isolate each session's product from all the others.

In answering question two ("What is the association between remediation and the diagnosis?") on the guide sheet, the subjects indicated that certain treatments were chosen because they were believed to be effective for certain weaknesses. Non-treatment statements fell into the "other" category as was the case for the first question.

Commonality

For all statements mentioned in two or more sessions, a given clinician's product for a given case included, on the average, roughly

half of the statements/associations mentioned two or more times for the same case by the other clinicians (Table 21). There was considerable variation among clinicians for a given product, with some sharing only a few statements or associations with group and others sharing more.

The key to putting commonality outcomes into proper perspective was the fact that, with the exception of cues, idiosyncratic statements (statements made during one session) for all products far outnumbered statements mentioned two or more times. Thus, similarity of performance, while appearing to be relatively high, was calculated on the basis of a restricted set of statements (those mentioned in two or more sessions). What was obscured was the fact that at most only one-third of a given product (excluding cues) was even eligible for inclusion in the commonality measurement. The commonality statistic indicated once again that only by aggregating products across sessions could the outlines of even a meager consensus on each case be revealed. The individual session products were not only largely idiosyncratic in content, they were also idiosyncratic in terms of format. There was no standard method used to present the information. Direct verbal comparison was thus virtually impossible.

Intercorrelations

According to the data shown in Table 22, the likelihood of any two clinicians agreeing on anything about a given case (excluding cues) decreased across products. At best, mean agreement was 14 percent of the combined statements (see Diagnosis column). At worst, mean agreement was one percent (see Remedial/Diagnostic column). The effect of

Table 21: Mean Commonality Scores for All Products

	Diagnosis	Remediation	Cues	Remediated Diagnosis	Remedial/Diagnostic Associations
Case 1	0.65 SD = 0.18	0.56 SD = 0.09	0.71 SD = 0.20	0.58 SD = 0.27	0.56 SD = 0.36
Case 2	0.53 SD = 0.23	0.59 SD = 0.11	0.65 SD = 0.27	0.52 SD = 0.30	0.38 SD = 0.26
Case 3	0.53 SD = 0.23	0.64 SD = 0.23	0.69 SD = 0.16	0.42 SD = 0.24	0.53 SD = 0.37
Case 4	0.66 SD = 0.11	0.60 SD = 0.08	0.65 SD = 0.22	0.19 SD = 0.08	0.45 SD = 0.19
Grand Mean	0.59	0.59	0.67	0.42	0.50

idiosyncratic statements (which was somewhat controlled for when products were aggregated across sessions) was highlighted in the results of the intercorrelations. A partial intercorrelation grid for the diagnoses of clinicians A and B is presented in Table 23. The clinician's agreed on seven percent of their combined statements.

Table 24 shows that the likelihood of clinicians agreeing with themselves on statements made about a case and its replicate (excluding cues) was only slightly more favorable than for two separate clinicians. At best, mean agreement was 23 percent of their combined statements (see Diagnostic column and Remedial column). At worst mean agreement, again, was one percent (Remedial/Diagnostic column). That is, they made wholly dissimilar statements during the two sessions. The impact of the idiosyncratic statements virtually overwhelmed the few points of agreement that did exist. A partial intracorrelation grid for the diagnoses of clinician A is presented in Table 25. The clinician agreed with herself on 20 (13) percent of the combined statements made in sessions one and three.

The sample matrices for inter- and intra-correlations illustrate sharply the very limited agreement within and between clinicians on a case. Individual differences in performance did exist, however, (Appendix P) with the widest variation shown for cues (0.07 to 0.77). However, wide individual differences in cues did not lead to equally wide ranges in performance on other products. Using the Porter statistic, the variability of diagnostic agreement and remedial agreement ranged from 0.00 to 0.29, of agreement on remediated diagnoses from 0.00 to 0.22, and on specific remedial diagnostic associations

Table 22: Mean Inter-Clinician Correlation

	Diagnosis	Remediation	Cues	Remediated Diagnosis	Remedial Diagnostic Associations
Case 1	.18 (.13)* SD=.10 SD=.06	.14 (.09) SD=.10 SD=.05	.24 (.35) SD=.26 SD=.18	.13 (.08) SD=.14 SD=.07	-.16 (.01) SD=.11
Case 2	.13 (.10) SD=.11 SD=.07	.17 (.13) SD=.09 SD=.08	.30 (.27) SD=.22 SD=.17	.13 (.08) SD=.13 SD=.07	-.18 (0.0) SD=.08
Case 3	.11 (.09) SD=.09 SD=.06	.10 (.07) SD=.10 SD=.05	.29 (.29) SD=.15 SD=.10	.04 (.04) SD=.09 SD=.05	-.18 (0.0) SD=.08
Case 4	.20 (.14) SD=.08 SD=.05	.15 (.11) SD=.10 SD=.06	.23 (.25) SD=.20 SD=.13	.20 (.13) SD=.08 SD=.05	-.18 (.01) SD=.07
Grand Mean	.16 (.11)	.14 (.10)	.24 (.29)	.13 (.08)	-.18 (0.0)

*Numbers in parentheses were calculated using the Porter statistic.

Table 23: Partial Intercorrelation Matrix for Diagnosis
Case 3, Clinicians A and B

Phi = .07
Porter = .07

STATEMENT ++	STATEMENT +-
Present in Both Clinicians A and B	Present in Clinician A, Absent in Clinician B
Hearing: Acuity (W) Speech articulation (W) Intellectual potential (S)	Vision: General (Obs) Vision: Acuity (W) Hearing: General (Obs) Affective: General (Obs) Emotional adjust (W) Parent education background (Obs) Parent attitude toward school (W) Parent-child relation (Obs) Sibling relationship (S) Comprehension: General (W) Comprehension: Oral (W) Comprehension: Silent (W) Comprehension: Vocabulary (W) Relationship with peers (W) Use of blends (W, Obs) Hearing acuity (Obs)
STATEMENT -+	STATEMENT --
Absent in Clinician A, Present in Clinician B	Absent in Both Clinicians A and B
Speech articulation (Obs) Motor coordination (W, Obs) Attitude toward reading instructional (W) Attitude toward reading independent (W) Ability to apply reading skills (W) Oral reading: General (Obs) Oral reading: Rate (W) Oral reading: Self correction (W) Oral reading: Phrasing (W) Oral reading: Intonation (W) Silent reading: General (W) Word analysis: General (W) Phonetic analysis: General (Obs) Use of initial consonant sounds (S) Use of syllables (W) Comprehension: Listening (S) Relationship with peers (W) Comprehension: Oral (S) Word recognition: General (S)	263 Domain statements excluded

Table 24: Mean Intra-Clinician Correlation

	Diagnosis	Remediation	Cues	Remediated Diagnosis	Remedial Diagnostic Associations
Case 1	.29 (.20)*	.34 (.20)	.66 (.65)	.30 (.18)	-.05 (.04)
Case 2	.12 (.09)	.33 (.23)	.40 (.41)	.14 (.09)	-.11 (.02)
Case 3	.23 (.17)	.17 (.12)	.38 (.34)	.18 (.10)	-.15 (.01)
Case 4	.27 (.19)	.33 (.23)	.22 (.27)	.26 (.17)	-.09 (.04)
Grand Mean	.23 (.14)	.29 (.20)	.31 (.49)	.22 (.14)	-.10 (.02)

*Numbers in parentheses were calculated using the Porter statistic.

Table 25: Partial Intracorrelation Matrix for Diagnosis
Case 3, Clinician A, Time One,
Clinician A, Time Two

Phi = .20
Porter = .13

STATEMENT ++ Present at Time One and Time Two	STATEMENT +- Present at Time One, Absent at Time Two
Motor coordination (W)	Attitude toward reading:
Intellectual potential:	Independent (obs)
General (S)	Motivation for reading (Obs)
Oral reading: Phrasing (W)	Emotional adjustment (W)
Oral reading: Intonation (W)	Substitutions contextually acceptable (W)
	Silent reading comprehension (W)
	Word analysis (W)
	Phonetic analysis (W)
	Comprehension vocabulary (Obs)
STATEMENT -+ Absent at Time One, Present at Time Two	STATEMENT -- Absent Both at Time One and Time Two
Motor coordination (Obs)	272 Domain statements excluded
Hearing acuity (W)	
Speech articulation (W, Obs)	
Attitude toward reading:	
Independent (W)	
Attitude toward reading:	
Instructional (W)	
Relationship to peers (W)	
Ability to apply reading skills (W)	
Oral reading: General (Obs)	
Oral reading: Rate (W)	
Oral reading:	
Self-correction (W)	
Silent reading: General (W)	
Word analysis: General (S)	
Phonetic analysis: General (Obs)	
Use of initial consonant sounds (S)	
Use of syllables (S)	
Word recognition: General (S)	
Comprehension: Oral (S)	
Comprehension: Listening (S)	

from 0.00 to 0.08. Clinicians who were highly reliable in cue collection, then, were not able to translate that into high reliability on other products. At best, the clinicians agreed (two or more sessions) on 29 percent of the total statements for the same case. At worst, on none of the statements. Therefore, the low mean inter-correlations shown in Tables 22 and 24 were not an artifact of some very high agreement levels being masked by low ones. Even at its best, intra- and inter-clinician reliability was markedly inadequate.

Comparison of Michigan and Illinois Subjects on All Products

The design of the study allowed for the comparison of the two groups of subjects: those who received their training entirely in Michigan and those who received their training entirely in Illinois.

Five t-tests for differences between mean performance on a case and its replicate (intra-clinician correlation) were computed to the test the hypothesis that there was no difference between the groups in their performance on the five products.

Using an alpha of .05, none of the five tests reached significance:

Diagnosis = 0.84

Remediation = 0.18

Cues = 0.07

Remediated Diagnoses = 0.38

Remediation/Diagnosis Associations = 0.31

There was no statistically significant difference in performance between subjects trained in Michigan and those trained in Illinois. However, the results were not uniform. The performance on cues

approached significance. The two groups appeared to have been trained differentially with respect to sensitivity to certain cue forms: lower inference forms (test scores and comments) and higher inference forms (test protocols and audio recordings). The performance on diagnostic statements, conversely, indicated that there was almost no difference whatever between the two groups on that dimension.

The Michigan subjects requested a total of 231 cues, 99 of which (0.42) were in the form of booklets and tapes. The Illinois subjects requested a total of 419 cues, only 38 of which (0.09) were in the form of booklets, and tapes. The two groups appeared to have been trained differently with respect to what was considered the most fruitful way to gather diagnostic information, but the differing methods did not translate into differing diagnostic performance. (Both groups examined Background Information to a similar extent: Michigan = 32 requests; Illinois = 44 requests.)

As had already been demonstrated, the most agreed upon cues (two or more sessions for a given case) were those that presented information as test scores and comments. This implied that, for the Michigan group, the examination of test booklets and audio tapes resulted in a high degree of idiosyncratic interpretation. For the Illinois group, the implication was that a large volume of information dealing with scores and comments was collected, also resulting in idiosyncratic interpretation.

Even though the Illinois subjects had a substantially higher level of intra-clinician agreement on which cues should be collected

(Illinois: $\bar{X} = 0.56$, Michigan: $\bar{X} = 0.27$) they still were not able to translate that agreement to diagnostic performance.

Correlational Data

(1) By Individual

The correlations were computed using mean intra-clinician figures compiled for each of the clinicians on each of the five products. The square of the value for the correlation (r^2), not the correlation coefficient per se (positive or negative) gives us information as to how well one variable can be predicted from knowledge of the other. Following are the correlational values computed for various sets of products.

(1) Cues and diagnosis ($r = 0.27$, $r^2 = 0.07$). We have seen that diagnoses were largely idiosyncratic, that large numbers of cues were collected, and that cues were interpreted in a variety of ways. Therefore, very little of a diagnosis could be predicted on the basis of knowledge of cues collected.

(2) Diagnosis and remediation ($r = 0.10$, $r^2 = 0.01$). Very simply, for a given clinician, on the average, the remediation could not be predicted on the basis of knowledge of the diagnosis. This refuted a major conclusion of the pilot study; namely that diagnosis and remediation at the clinician level were moderately correlated.

(3) Cues and remediation ($r = 0.25$, $r^2 = 0.06$). As with diagnosis, very little of the remediation could be predicted on the basis of knowledge of cues collected.

(4) Remediated diagnosis and remediation ($r = 0.25$, $r^2 = 0.002$).

Knowledge of the diagnostic statements that were designated as requiring remediation provided no predictive power as to the specific remediations that would be suggested when performance was averaged across two sessions for the same case.

(2) By Group

The correlations were computed using mean commonality figures compiled for each case (Table 22). Commonality scores were used to approximate group performance on all products in order to see whether group performance was more predictable than individual performance. Predictive capability was found to be low even at this aggregate level.

1. Cues and diagnosis: ($r = 0.14$, $r^2 = 0.02$)
2. Diagnosis and remediation: ($r = -0.58$, $r^2 = 0.33$)
3. Cues and remediation: ($r = -0.18$, $r^2 = 0.03$)
4. Remediated diagnosis and remediation: ($r = -0.43$, $r^2 = 0.18$)

Again, knowledge of cues collected provided no grounds for predicting either diagnosis or remediation. However, some of the variation in agreed upon remediation could be accounted for by knowledge of the diagnosis and the remediated diagnosis.

Clinician Self-Reports

At the close of the third session, each clinician was asked to self-report on how she normally went about gathering information, writing a diagnosis, and relating the initial remediation plan to the diagnosis. They reported collecting many of the same types of information, e.g., on intelligence scores, comprehension, word attack and

recognition skills, etc. Further, they shared very similar procedures for writing their diagnosis. They said that they organized their diagnosis (1) in the order in which they collected information from the children, or (2) by grouping their findings by strengths and weaknesses, or (3) by listing findings according to areas of need, together with supporting information. Finally, they held similar views as to the remediation. They asserted that (1) there should be a direct one-to-one relationship between the two, although there will be some things that cannot be remediated; (2) weaknesses should be treated by using strengths; and (3) weaknesses should be prescribed for directly.

The subjects' responses to all three questions were prompt and very similar, suggesting that they had internalized a standard rhetoric for describing the task of diagnosis and remediation.

Following is an example of clinician C's stated plan. The mean intracorrelation for this clinician's diagnosis was .18 which was slightly above the group mean. Eighteen percent of the statements made for the case at time one were repeated again at time two.

#1: What guides you in determining what information to collect and then what guides you in writing up the diagnosis?

First I want to get a general impression of how he's doing, first in comprehension. If he's doing poorly in comprehension, I try to find some reasons for it. First I look at word recognition and then I look at phonics. Then I try to group the test scores that have to do with word recognition and those that have to do with phonics. If two phonics scores didn't seem to match I would look at the test to see what exactly the kid had to do, what the task was. If one score was much lower than another, what was it about the first test that was so much harder for the kid?

If the comprehension is low I usually look for an I.Q. score or a listening score to see if they have the potential for higher comprehension. If a kid had a high I.Q. and they were doing poorly in reading, then I would look at hearing, talk to the parents, ask for physical stuff. Right now I don't go into a kid's emotional state, for example, unless it's very obvious, like the kid is aggressive or very shy. I just try to deal with the skills themselves.

In the write-up I listed an observation and then the scores that supported it. First I look at comprehension, then vocabulary, then phonics. But if there was something in the scores that didn't seem pertinent to the diagnosis then I just left it out. Usually I try to identify just weaknesses. An observation that he did well in arithmetic doesn't make a big difference so I just left it out of the write-up.

#2: What do you see as being the relationship between a diagnosis and initial remediation plan?

It should be one to one. If you notice that something is a weakness then you should write steps to overcome this weakness. If there's a strength in reading you should try and use that strength and maybe apply some of them. You wouldn't give a kid exercises to practice something he knows, though. This kid seemed to be better in sight words than in comprehension but he wasn't very good on suffixes so I figured that if he had sight words that contained some suffixes you could say, "look at the suffix." So the remediation should be based almost completely on the diagnosis. But once you start the remediation if it doesn't work then you have to revise it. Working with the kid you find more than what the test scores tell you.

If clinician C followed her plan, she would collect information in the following order: comprehension, word recognition, phonics, I.Q., and physical information. According to the plan, the diagnostic write-ups would present the findings in the order indicated. Finally, there would be a one-to-one relationship between weaknesses stated and remediations offered.

The actual performance of the clinician was at variance with the stated plan. With respect to cue collection, information was collected on comprehension, word recognition, phonics, I.Q., and physical

factors. However, the order was different for each of the three cases. In addition, considerable additional information not mentioned in the plan was collected including: background information, arithmetic and spelling samples, informal oral reading, individual achievement tests, and writing samples. There was no discernable pattern of collection for this additional information.

Across the eight clinicians, two overall observations were made regarding cue collection.

1. The order in which cues were actually collected across sessions did not follow the order given in the stated plan. Further, whatever order was used, it was never followed a second time, either for the middle (distractor) case or for the replicate case.
2. A substantial amount of information was collected that was never mentioned in the stated plan at all. Six of the eight clinicians requested this additional information to the extent that it equalled or exceeded the information that was collected according to plan.

While it was possible that this information might have been requested merely because it was available, the data warrant the interpretation that the clinicians (1) had no ready framework within which to fit the new data, and (2) had no commonly shared mode of interpreting what they did request.

In sum, the data indicated that the cues lead the clinicians, not the other way around. Without an explicit heuristic to guide data collection, each clinician essentially was at the mercy of whatever

information she chose to collect. The resulting information load had to be simplified and made manageable, but again, in the absence of an agreed upon model of consistent data collection and interpretation, that process became idiosyncratic and therefore unpredictable across sessions.

With respect to writing the diagnosis, the statements in the sample were not ordered in any consistent way, whether comparing the diagnosis for the case and its replicate or comparing either diagnosis with the plan. Following are the first paragraphs from clinician C's diagnostic write-ups for the case and its replicate (sessions one and three). Clinician C did follow that stated procedure of supporting observations with scores.

Session 1. "Poor auditory acuity may be an inhibiting factor in his reading. His test in auditory acuity was below normal. He missed 5 of the 30 pairs on the Wepman. He scored only 2.5 grade level on the 'Hearing Sounds in Words' test on the Durrell."

Session 3. "His verbal skills seem to be below average for his age. His verbal WISC score was 86, and his listening comprehension was 4.5. He does better in areas not requiring reading. His performance I.Q. was 106."

Across the eight clinicians, three overall observations were made regarding diagnostic write-ups.

1. In no instance did the diagnostic statements follow the order of cues collected.

2. Whatever the order in which diagnostic statements were presented, no clinician repeated that order a second time.
3. In those cases where strengths and weaknesses were loosely grouped and supporting information was provided, the pattern was never the same for all three sessions for a clinician, and the order of presentation always varied.

Finally, clinician C's remedial suggestions were not presented in the same order as diagnostic statements. Further, in addition to one-to-one associations, there were often multiple diagnostic statements associated with a given remediation. A sample of some of clinician C's associations follow.

Remediation	Diagnosis
<u>Session 1</u>	
To increase comprehension level use exercises such as the cloze or modified cloze. (T)	His comprehension scores are lower than his sight words. (W)
<u>Session 3</u>	
Ask questions after each selection to train him to attend to meaning. (T)	His comprehension was only 71 percent. (W)
	In most cases his errors resulted in loss of meaning. (W)
	His silent reading scores are very low. (W)
	His rate is low. (W)

Across the eight clinicians, three overall observations were made regarding the associations made between remediation and diagnosis.

1. In no instance did the remedial statements follow the order of presentation of the diagnosis. Typically, remedial statement number one, for example, might be paired with diagnostic statement numbers 12, 13, and 24.
2. There was no evidence of systematic use of weaknesses being treated by using strengths.
3. No clinician had a one-to-one match between diagnosis and remediation. There were always unmatched diagnostic statements.

Comments About the Sessions

The subjects' responses were highly positive. Almost all clinicians characterized the sessions as being both enjoyable and educational. Several hoped that the study might have an impact on the design of training of diagnosticians. Others felt that the sessions had sharpened their understanding of what they were doing.

CHAPTER FIVE

DISCUSSION

The assumption in the literature that diagnosis and remediation are, or ought to be, routinely and reliably associated is not remotely substantiated by any of the findings of this study except at the case level. Even at that aggregate level, the association is only modest.

Across SIMCASES clinicians examined information about background and cognitive ability and the results of performance on diagnostic tests, graded word lists, reading achievement tests, basic sight vocabulary and visual/auditory tests. Thus they all began with essentially the same categories of information for all the cases (although the order of collection was entirely unpredictable). From this common information base they went on to generate diagnoses and remediations. Only one-quarter to one-third of the statements comprising these two products were mentioned in two or more sessions. The proportion of agreed-upon statements fell precipitously when the standard for examining reliability changed to agreement in three or more sessions.

As the data reported has shown, there was little association between diagnosis and remediation. A telling example could be seen in Case 3. There were 17 diagnostic statements which were agreed upon as characterizing the case but only six remediations were agreed upon as constituting an appropriate action plan. Two of those were unrelated to any of the agreed-upon diagnoses (Table 6, Table 9).

Additionally, with the exception of Case 1, the group diagnoses were at variance with the "ideal" diagnosis (Appendix 0) compiled by the senior reading clinician: omission of weaknesses in fluent texting and comprehension in Case 2; omission of weaknesses in sight vocabulary and phonetic analysis in Case 3; and omission of weaknesses in multisyllabic analysis, fluent texting, and auditory problems in Case 4.

Thus, the clinicians may have agreed as a group that certain problems characterized a case but, as a group, they may have been largely incorrect in their judgments. Conversely, they may have compiled a correct group diagnosis and the judgments of the senior clinician may be in error. The question of the validity of the diagnoses (as distinguished from their reliability--one can be reliably wrong) was beyond the scope of this investigation. However, it is clearly a major unresolved issue and is discussed further in Chapter Six.

What the data in this study unambiguously demonstrated was that the clinicians studied were not reliable diagnosticians. They believed they were acting consistently but in fact collected a great deal of information in unpredictable fashion, most of which was interpreted idiosyncratically.

Why did the clinicians in this study behave as they did?

Speculation #1: task and external validity contributed to low reliability. The subjects' performance might have been an artifact of the experimental setting, including the use of SIMCASES. That is, the task itself was crafted so that while the clinicians could perform comfortably as required, they weren't doing what they actually do on

the job. Perhaps a naturalistic setting using real children would have yielded different results. However, any unreliability demonstrated in a naturalistic setting could be attributed to the differing behaviors of the children themselves. Both settings contribute to problems of task validity. Nevertheless, a major assumption in our observation studies as well as in training and research studies in medicine, is that important problem-solving behaviors of clinicians can be elicited through simulated cases. "At least some of the important cognitive skills are elicited by sets of data gathered about real cases with real problems and stored for presentation as simulated cases" (Vinsonhaler, 1979).

The question of external validity, or generalizability to a larger population, is complicated by the small sample size. However, a major trade-off for small sample size is the power of replication. As discussed earlier, a number of studies of experienced reading specialists conducted under the same conditions as the present study and using the same SIMCASES reported diagnostic results almost identical to those in this study (Weinshank, 1980). This is the first study to investigate remedial reliability and that of remedial/diagnostic associations, but there is no reason, in principle, to think that these results would not be replicated with another sample.

A second issue in external validity is the nature of the sample itself. Was this a representative sample of experienced, credentialled reading specialists or did chance dictate selection of an unrepresentatively incompetent group? These subjects met or exceeded the criteria set by the study. They were all employed in good standing

in five different school districts in two states. They were involved in professional reading organizations and/or had been recommended by their academic mentors as being highly competent students. If anything, one might argue that they were an unrepresentatively able group, thus making the results all the more disquieting.

Speculation #2: training influenced performance. Contemporary teacher training programs are committed to a position that calls for attending to individual differences ("meeting individual needs"). This view is not compatible with the proposition that a small number of diagnostic and remedial categories account for a large proportion of the problems exhibited by deficient readers of a given age and grade range. Training that embodies the individual differences paradigm has no need of predictive capability. By definition, if everyone is different in significant ways that must be uncovered, student A's performance cannot and should not have any relationship to student B's performance. If each student's reading problem is approached from the standpoint that the problem is different in significant ways from everyone else's, those differences must somehow be divined from the cues. Which cues are the key ones? There is no evidence in this study that the training process has addressed itself to this question, as evidenced by the almost complete lack of diagnostic and remedial agreement between clinicians and by each clinician with herself on the same case.

If the individual differences hypothesis is correct, then the clinicians in this study are performing precisely as they have been trained. Making between 20 and 30 diagnostic statements and between

25 and 35 remedial statements per session (as almost all did) would surely lead a clinician to believe that individual needs had been taken fully into consideration. When coupled with unsystematic cue collection, the results of any cross-session comparison would be almost guaranteed to show no association, even when the case being examined is the replicate of an earlier case.

Speculation #3: the clinical setting itself does not provide a context for the specialist's learning from experience. Diagnostic judgments are reached and remedial suggestions made, but there is no way to get feedback about the consequences of recommended actions. The clinician simply does what the profession dictates with no opportunity to confirm or disconfirm hypotheses about problems and their treatment.

Speculation #4: it is in the nature of the profession to mystify the process so that it merits the respect it claims for itself. Perhaps the painful reality is that most diagnoses are embarrassingly inflated and clinicians know that such diagnoses ought not to be taken seriously when it comes to remediation. Similarly, clinicians may understand all too well the limited range of action alternatives open to them but professional mystique demands that they once again engage in clinical overkill.

Speculation #5: the clinicians had a plan but they didn't have a satisfactory method. According to Newell and Simon (1972), "A method is simply a plan that you use twice" (p. 835). The plans they believed they were following to organize their behavior were never actually

carried out twice in the same way; there was no way to predict behavior from case to case. The method actually used may be the answer to the question, "How can you make a straightforward instructional problem look very complicated?" Answer: disguise the fact that, in general, a very limited number of diagnostic and remedial judgments reliably characterize most cases of reading difficulty. Produce a verbal jungle of strengths, observations, and weaknesses derived from a needlessly exhaustive examination of case information collected unsystematically. This procedure is guaranteed to take quite a bit of time, look substantive, and impart a feeling of professionalism.

Speculation #6: the clinicians have not mapped onto the task any model of the process of acquisition and mastery of reading behaviors. Lacking such a model, their search for and interpretation of information is unordered and the resulting instructional strategies are unsequenced.

None of these speculations is flattering to the field of remedial reading. But the results of this study are such that unless one were to assert that the subjects were "lemons" and the experimental conditions spurious (charges that have been examined and refuted), the near-chaotic state of diagnosis and remediation cannot go unchallenged.

CHAPTER SIX

SUMMARY AND RECOMMENDATIONS

Introduction

The purpose of this study was to establish the nature and extent of the relationship between diagnosis and remediation in reading. The study empirically investigated the clinical problem solving behavior of eight reading specialists as they (1) diagnosed simulated cases of reading difficulty, (2) prepared an initial remediation plan, and (3) associated remedial statements with diagnostic statements.

Summary of Procedures

The eight subjects in the study were practicing, experienced, credentialled reading specialists. Four received their training entirely in Michigan and four entirely in Illinois. Each subject participated in three sessions in a controlled setting, each session lasting from two and one-half to three hours. Subjects were randomly assigned to simulated cases (SIMCASES) of reading difficulty. Each simulated case was based on a real child who was served by the Reading Clinic at Michigan State University.

In a total of 24 sessions they (1) wrote a diagnosis for a SIMCASE of reading difficulty; (2) wrote an initial treatment plan; (3) transferred diagnostic statements to a standardized diagnostic checklist and remedial statements to a standardized remedial checklist;

(4) made associations between remedial and diagnostic statements; (5) reported on reasons for making the associations; (6) reported on reasons for the presence of unassociated diagnostic elements (if any) and unassociated remedial elements (if any); and (7) answered three free-response questions.

The products that resulted from these procedures were:

1. Diagnostic statements;
2. Remedial statements;
3. Cues;
4. Remediated diagnoses; and
5. Specific remedial/diagnostic associations.

Summary of Results

The products that resulted from each session were analyzed to determine (1) the extent of agreement among and between clinicians (proportional agreement, commonality, intercorrelations; (2) strength of correlations between products; and (3) extent of differences in performance between Illinois-trained and Michigan-trained subjects.

Proportional Agreement

A very small number of statements/associations accounted for agreement across two or more sessions for a given case. The bulk of each product was composed of idiosyncratic statements, i.e., statements made in only one session.

Both redundant diagnostic and remediated diagnostic statements were more likely to be characterized as a weakness than as a strength or observation.

Redundant remedial statements were more likely to be characterized as a treatment than as a strength, weakness, or observation.

Redundant associations of remedial and diagnostic statements were more likely to pair treatment with weakness than with any other category.

Across all cases, a limited number of categories of diagnostic and remedial statements accounted for all statements made two or more times.

There was substantially higher agreement for cues requested across sessions than for the other products. Cues requested two or more times were much more likely to be in the lower inference form of test scores and comments than in the higher inference form of test booklets and audio tapes. Examination of common cues leads neither to common diagnoses, common remediation, nor common associations between remediation and diagnosis.

Commonality

The results of comparing an individual's product with those of the group (commonality) were consistent across all products. Considering only statements mentioned two or more times, a given clinician's product for a given case included on the average, roughly half of the statements/associations mentioned during the other sessions. While this may appear to be a relatively high level of agreement between individual and group, it must be remembered that commonality was calculated on the basis of a restricted set of statements, i.e., those mentioned in two or more sessions. At most, only one-third of any product (excluding cues) would be included in the calculation.

Both the proportional agreement and commonality results indicated that only by aggregating products across sessions could the outlines of a meager consensus on each case be demonstrated.

Intercorrelations

With the exception of cues, the results across products followed a consistent pattern. The likelihood of any two clinicians, or one clinician over two sessions, agreeing on a statement/association for a given case ranged from small to non-existent.

Comparison of Subjects

Five t-tests for differences between means were computed to test the hypothesis that there was no difference between the Michigan and Illinois subjects in their performance on the five products. None of the tests reached statistical significance at the .05 level but the difference in cues approached significance. Michigan subjects requested fewer cues and proportionally more high inference cue forms than did the Illinois subject (0.42 vs. 0.09). Nevertheless, the differing methods did not translate into differing diagnostic performance. Statistically, the groups were virtually identical in performance on that dimension indicating that irrespective of number or form of cues collected, there was no agreed upon heuristic to guide information gathering and interpretation.

Correlational Data: Individual and Group

At the individual clinician level, performance on one product could not be used to predict performance on another when performance was averaged across two sessions for the case and its replicate. Cues

were predictors neither for the diagnosis nor for the remediation. Neither was the diagnosis a predictor for the remediation.

At the group level, only diagnosis and remediation showed a modest level of association. However, cues did not predict either diagnosis or remediation even at this aggregated level.

Clinician Self-Reports

Each clinician reported on plans of action that she used to guide her information gathering and the writing of her diagnosis and initial remediation. An examination of teacher performance on three products (cues, written diagnoses, written remediations) revealed that the subjects never followed their stated plans.

- a. Cues never followed a predictable order of collection.
- b. Diagnostic statements never followed the order of cue collection. Whatever the order in which diagnostic statements were presented, no clinician repeated that order a second time.
- c. Remedial statements never followed the order of presentation of diagnostic statements.

In sum, no clinician followed a model or heuristic powerful enough to allow for predictable collection of data and its consistent interpretation.

Conclusions

Two major hypotheses which guided the design of this study were weakly supported by the data.

1. Statements of weakness were repeated more reliably across sessions for a given case than were all diagnostic categories.
2. Statements of weakness were more often associated with treatments across sessions.

Two additional hypotheses dealing with the nature of unassociated diagnostic and remedial statements clearly were strongly supported by the data.

1. Those statements in the diagnosis that were not associated with any treatments remained unassociated because:
 - a. they could not be remediated,
 - b. they did not require remediation,
 - c. the clinician meant to provide a remediation but forgot to write it down, and
 - d. the clinician knew no remediations for those diagnoses.
2. Those statements in the remediation that were not associated with any diagnosis remained unassociated because the clinician believed that the particular remediations should always be included.

A final hypothesis, that remediated diagnosis would be repeated more reliably across sessions than all diagnostic categories, was not supported. There is almost no agreement on this dimension.

Assuming that the subjects in this study were representative of a larger group of similarly qualified and experienced reading specialists, what can be concluded from the data with regard to the efficacy of diagnostic procedures?

First, as currently practiced, the compilation and presentation of diagnostic findings is an empty, random exercise. There are as many different conclusions for a case as there are points of entry into the data. All statements beyond those few which fall within the seven general diagnostic categories are entirely idiosyncratic, hence, unpredictable.

Second, all remedial suggestions beyond those few reflected by the seven general remedial categories are idiosyncratic and may or may not (as chance dictates) have any impact on the child's performance.

Given the very low level of diagnostic predictability demonstrated in this study, it can be argued that diagnostically derived remediation is neither time efficient nor economically justifiable. Any combination of remediations chosen from the seven categories might do equally well in addressing the actual reading problem(s) of a student.

Recommendations

A number of options present themselves as policy alternatives.

(1) The case for the ultimate desirability of the current diagnosis-treatment model could continue to be pressed. The data in this study unambiguously refuted the effectiveness of this approach. A plan of instruction based on an ad hoc diagnosis cannot be defended. It speaks neither to the child's legitimate educational needs nor to the issues of the time and cost of preparing the diagnosis to begin with.

(2) Diagnosis as a necessary prerequisite to remediation could be eliminated in favor of simply choosing those remediations that work

most of the time for most children at a given level of reading proficiency. Unfortunately, we have no long-term empirical data describing the problems that characterize most problem readers at a given age and grade level nor the remediations of choice for those problems. We are, therefore, still left with less than optimal control over which remediation or combination of remediations is the best starting point for a student.

(3) A reading-performance diagnostic model could be developed that concerns itself first and foremost with evaluating reading performance against a commonly accepted metric which would include, in a standard sequence, the tasks embedded in the diagnostic categories. Instead of collecting information for clues that might shed some light on the particular difficulties that characterize an individual student, the procedure would be reversed. The reading-performance diagnostic model would dictate what information should be collected and in what sequence. By assuming that most children's deficiencies can be grouped within a small number of categories, cue collection becomes routinized and initial remediation follows directly from the information collected. Those anomalous children who do not respond (within specified time boundaries) to standard remediation(s) would be re-evaluated.

Which of the three models would lead to significantly better outcomes in improved reading proficiency? This is an empirical question which is beyond the scope of this investigation. Nevertheless, it can be suggested that the use of controlled clinical trials, so common in other disciplines, could be established. University

clinics could be used as sites for such long-term investigation, with incoming students being randomly assigned to remediation:

- a. that derives from the traditional two or three hour diagnostic work-up, or
- b. that is not preceded by any individualized diagnosis but that progressively clarifies the problem, thereby producing the diagnosis, or
- c. that derives from a reading-performance model of diagnosis.

Outcome data would be collected by people blind to the experimental condition of a given child.

In "The Structure of Scientific Revolutions," (in Neurath, ed.)

Thomas Kuhn says:

...the emergence of new theories is generally preceded by a period of pronounced professional insecurity...That insecurity is generated by the persistent failure of the puzzles of normal science to come out as they should. Failure of existing rules is the prelude to a search for new ones.

The paradigm in the field of reading diagnosis and remediation has been that diagnosis and remediation ought to be integrally related. The major conclusion of this study is that they are not.

Kuhn says:

Just because the emergence of a new theory breaks with one tradition of scientific practice and introduces a new one conducted under different rules and within a different universe of discourse, it is likely to occur when the first tradition is felt to have gone badly astray.

The results of this study directly challenge the reigning paradigm in the field of reading diagnosis/remediation. If these results are substantiated, they would invalidate a large part of

the training of reading diagnosticians together with the textbooks which they study.

"Paradigm-testing," Kuhn argues, "occurs only after persistent failure to solve a noteworthy puzzle has given rise to crisis." In this thesis, the puzzle is that the reading diagnosis/remediation paradigm does not give rise to any generalizable predictive metric exemplified in the work of the "community of practitioners" (Kuhn's description) among whom are the reading clinicians in this study.

APPENDICES

APPENDIX A
CUE INVENTORY - CASE 4

APPENDIX A

CUE INVENTORY - CASE 4

Case Information Inventory

Types of Information	Forms of Information			
	Test Scores	Examiner's Comments	Test Booklet	Audio Recording of Session
BACKGROUND INFORMATION				
Teacher Form			BKG9	
Parent Form			BKG3	
School Record			BKG15	
School Information			BKG19	
STANDARDIZED ASSESSMENT				
Cognitive Ability (Wechsler Intelligence Scale for Children Revised)				
Verbal Scale	WISR1			WISR5
Performance Scale	WISR7			WISR11
Full Scale	WISR13	WISR14	WISR15	
Graded Word List (Slosson Oral Reading Test)	SORT1	SORT2	SORT3	SORT5

Types of Information	Forms of Information				Test Directions
	Test Scores	Examiner's Comments	Test Booklet	Audio Recording of Session	
STANDARDIZED ASSESSMENT (Continued)					
Durrell Listening-Reading Series, Intermediate Level Vocabulary Paragraphs	DLR1 DLR7	DLR2 DLR8	DLR3 DLR9		DLR5
Achievement Test (Iowa Test of Basic Skills) Vocabulary Subtest Reading Subtest	ITBS1 ITBS7	ITBS2 ITBS8	ITBS3 ITBS9		ITBS5 ITBS5
Reading Achievement (Gates Macginitie) Speed/Accuracy Test	GMG1		GMG3		GMG5
Diagnostic Analysis of Reading Difficulty (Durrell) Oral Reading Silent Reading Listening Comprehension Word Recognition and Word Analysis Hearing Sounds in Words-Primary Visual Memory of Words-Primary Intermediate Spell, List 1	DUR1 DUR7 DUR13 DUR19 DUR25 DUR31 DUR37	DUR2 DUR8 DUR14 DUR20 DUR26 DUR32 DUR38	DUR3 DUR9 DUR15 DUR21 DUR27 DUR33 DUR39	DUR4 DUR10 DUR22	DUR5 DUR11 DUR17 DUR23 DUR29 DUR35 DUR41

Types of Information	Forms of Information			
	Test Scores	Examiner's Comments	Test Booklet	Audio Recording of Session Test Directions
STANDARDIZED ASSESSMENT (Continued)				
Phonic Spelling of Words	DUR43	DUR44	DUR45	DUR47
PHYSICAL INFORMATION				
Audiometric Record Vision Test		AUD2 VIS2	AUD3 VIS3	
NON-STANDARDIZED ASSESSMENT				
Auditory Discrimination (Wepman Auditory Discrim. Test)	WEP1	WEP2	WEP3	WEP5
Gates McKillop Reading Diagnostic Tests				
Recog. and Blending				
Common Word Parts	GMK1	GMK2	GMK3	GMK4
Auditory Blending	GMK7	GMK8	GMK9	GMK10
Giving Letter Sounds	GMK13	GMK14	GMK15	GMK16
SENTENCE COMPLETION				
Basic Sight Vocabulary (Dolch Word List)	DOL1	DOL2	DOL3	DOL5

APPENDIX B
OBSERVATIONAL STUDY INSTRUCTIONS

APPENDIX B
OBSERVATIONAL STUDY INSTRUCTIONS

You have been asked to serve as a consultant for the Institute for Research on Teaching at Michigan State University in order that we may better understand how people like yourself diagnose and remediate for children with reading problems.

Because this investigation takes several hours, the IRT will pay you as a consultant for the College of Education. A check will be mailed to you when your consulting work is completed. Please fill out this consultant pay form. (Give clinician pay form.)

The Institute is required by law to protect your privacy by keeping confidential your name and social security number. Your name will not be part of Institute permanent records. Instead, a number will be used as identification.

Please read this Informed Consent Sheet. Your signature indicates that you agree to participate in the study and allow us to use the resulting information. (Give time to read and sign Consent Sheet.)

Please take a few minutes, now, to fill out this background information sheet. (Give clinician Background Information Sheet.)

I would now like to give you an overview of the three sessions we will have together. In all sessions you will do two things. First, you will examine a simulated case of reading difficulty and will write a diagnosis and initial remediation plan. Second, you will focus on

specific aspects of the written diagnosis and remediation. Finally, in the third session only, you will be asked to reflect on some of your decisions.

Before we begin, we will practice using a simulated case of reading difficulty. Here is a file box in which is stored information about a child with reading difficulties. The information is available to you in five possible forms as shown on this cue inventory (give clinician copy of inventory): (1) test scores; (2) examiner's comments; (3) test booklets; (4) audio tapes; and (5) test directions. (Point to each category as it is being read.)

The cue inventory for this practice case tells you specifically what information is available within each form. For example, the Durrell Oral Reading Test is available to you in all five forms (show example of each form as it is being read): as a test score; examiner's comments; test booklet; audio tape; and test directions.

You request information by referring to the inventory keyword. For example, if you want the results of the Durrell Oral Reading Test in the form of examiner's comments, you would get it by asking for the keyword DUR----. I would then hand you that piece of information. You may request one item of information at a time.

This cue inventory is merely a listing of information that has already been collected for you. When you request specific forms of information do so in the order you normally follow when working with a child.

You may now request two items of information using the inventory keywords. I will hand you items you request. (Give the clinician time to make the requests.)

Having practiced using a case, you are now ready to request information about the case you will use this session. You will use this information to determine the most likely diagnosis and to suggest an initial remediation plan. Do not feel you must request an item of information just because it is present in the inventory; there is no right or wrong amount of information to request.

You may take notes if you wish and you may keep all forms of information requested until you have completed your work with the case.

After the presentation of some initial information about the case, you will have 30 minutes to collect information and reach a decision about the diagnosis. You will then have up to 20 minutes in which to write down your diagnosis and up to 20 minutes in which to write the initial remediation plan.

To review:

1. Collect information in the order you normally follow, using the keywords;
2. Take notes if you wish;
3. Keep the information requested if you wish;
4. Take up to 30 minutes to reach a decision about the diagnosis.

I will let you know when 10 minutes remain. Do you have any questions? (Pause)

You are now ready to begin examining a simulated case of reading difficulty. The case you are to consider today concerns a ____ year old boy named _____. Here is a sketch of the child, a statement

of the reason for referral to the reading specialist, and a taped interview which I will now play for you. (Play tape.)

I will start timing for 30 minutes. You may now request items of information.

BEGIN TAPE RECORDER

ENTER TIME/CUES ON CUE COLLECTION FORM

STOP AFTER 30 MINUTES

Would you now write your diagnosis? Please write complete sentences, not just key words or phrases. Assume that the report will be read by a clinician with training similar to yours. You will have up to 20 minutes to write your diagnosis. I will let you know when five minutes remain. (Provide clinician with wide-lined paper.)

Would you now write your initial remediation plan? Please write complete sentences, not just key words or phrases. Assume the plan will be used by a clinician with training similar to yours who will be responsible for working with the child. You will have up to 20 minutes to write your remediation plan. I will let you know when five minutes remain. Do you wish to keep the cues you've collected? (Provide clinician with wide-lined paper.)

BREAK

We are now going to begin the second part of our session during which you will focus on specific aspects of the written diagnosis and remediation.

The first step is for you to identify all of your key diagnostic elements. Here are a few statements from an abbreviated sample diagnosis. (Give sample diagnosis.) This diagnosis is not meant to

suggest a model; it merely provides practice in identifying types of diagnostic elements. Notice that the clinician circled what he or she considered to be key diagnostic elements and then numbered each circled element in turn. Notice, too, that the circle was left open at the end of the line when the element continued down to the next line. (Point out circled, numbered elements.)

To complete the process, the clinician coded each diagnostic element as being either a strength, weakness, or observation, by placing an S for strength, W for weakness, or Obs for observation next to each circled numbered element. (Point to codes.) In contrast to a strength or weakness, an observation is simply a neutral statement that characterizes some aspect of the case.

I would now like you to practice identifying key diagnostic elements. First circle and number, what you consider to be the remaining diagnostic elements in the sample diagnosis, and then code each one. (Give time to complete "Practice" section.)

I would now like you to do this task with your own written diagnosis. The sample diagnosis in no way implies that you should have a certain number of each type of element. It was designed merely to provide practice in identifying types of elements.

Use this red pen to circle, number, and code the diagnostic elements. (Circle, number, and code own diagnosis.)

You will now match the circled elements to statements on this diagnostic checklist. Here is a sample checklist. (Give Diagnostic Checklist.) The checklist is divided into 12 major categories which appear as capitalized headings. (Point to categories and give

clinician the separate table of contents sheet.) This table of contents sheet will aid you in rapidly locating the categories.

The clinician who wrote the sample diagnosis decided that diagnostic element Number One fell under the main category FAMILY. (Point to appropriate heading.) Under this category, the clinician further decided that element Number One was most similar to the sub-category "Sibling Relationships." Since it had already been coded Obs for Observation, the clinician merely placed the number one under the appropriate column heading, in this case Observation. [Point to the number (1) in the checklist.] Please observe how the clinician transferred diagnostic elements 2 and 3. (Give time to look at sample diagnosis and checklist.)

Would you like to review the steps we've just gone through?

I would like you to practice matching your circled, numbered, and coded practice diagnostic elements to diagnostic elements in the sample checklist. (Give time to match.)

Here is another copy of the Diagnostic Checklist and your written diagnosis. Please match your written diagnostic elements to the checklist according to the steps we just reviewed. If you have a diagnostic element that does not correspond to any of the main categories, list it under the main category OTHER. (Point out.) If you have a diagnostic element that corresponds to a main category but does not correspond to any sub-category under it, list it under the appropriate "Other" category. (Point out.) Please make as limited use as possible of these "Other" categories.

You will not be timed. Please work as rapidly as is comfortable. You may begin.

(When completed, make sure all the written diagnostic numbers are accounted for in the Diagnostic Checklist.) Read off numbers and corresponding codes from checklist. Clinician reads along in his/her written diagnosis.

Please turn your attention now to your remediation plan. You will be asked to identify all of your key remedial elements. Here are a few statements from an abbreviated sample remediation. This remediation is not meant to suggest a model. It merely provides practice in identifying types of remedial elements. Notice that the clinician circled what she or he considered to be the key remedial elements and then numbered each circled element in turn. Notice that the circle was left open at the end of the line whenever the element continued down to the next line. (Point out circled, numbered elements.) To complete the process, the clinician coded each remedial element as being either a strength, weakness, observation, or treatment, by placing an S for strength, W for weakness, Obs for observation, or T for treatment next to each circled, numbered element. (Point to coded elements.)

I would now like you to practice identifying key remedial elements. First, circle and number, and then code what you consider to be the remaining remedial elements in the sample remediation. (Give time to complete "practice" section of remediation.)

I would now like you to do this task with your own written remediation. The sample remediation in no way implies that you

should have a certain number of each type of element. It is designed solely to provide practice in identifying types of elements.

Use this red pen to circle, number and code the remedial elements. (Circle, number, and code own remediation.)

You will now match the circled elements to statements on this remedial checklist. Here is a sample checklist. (Give remedial checklist.)

This sample checklist is divided into 12 major categories which appear as capitalized headings. (Point to categories and give clinician the separate table of contents sheet.) This table of contents sheet will aid you in rapidly locating the categories.

The clinician who wrote the sample remediation decided that remedial element Number One fell under the main category VISUAL. (Point to appropriate heading.) Under this main category, the clinician further decided that element Number One was most similar to the sub-category Discriminate Visually Similar Words. Since it had already been coded T for treatment, the clinician merely placed the number (1) under the appropriate column heading, in this case T. [Point to the number (1) in the checklist.] In this particular instance, there was more in the written element than was present on the checklist. The clinician wrote in the additional comment, "Highlight the middle." (Point to written-in comment.) Please notice how the clinician transferred remedial elements 2, 3, and 4. (Give time to look at sample remediation and checklist.)

Would you like to review the steps we've just gone through?

I would like you to practice matching your circled, numbered, and coded practice remedial elements to remedial elements in the sample checklist. (Give time to match.)

Here is another copy of the Remedial Checklist and your written remediation. I would like for you to match your written remedial elements to the checklist according to the steps we just reviewed. (Point.) Remember: You can add additional remedial comments in the space provided. (Point.) If you have a remedial element that does not correspond to any of the main categories, list it under the main category labelled OTHER. (Point out.) If you have a remedial element that corresponds to a main category but does not correspond to any sub-category under it, list it under the appropriate "Other" category. (Point out.) Please make as limited use as possible of these "Other" categories. You will not be timed. Please work as rapidly as is comfortable. (Read off numbers and corresponding codes from checklist. Clinician follows along in his/her written remediation.)

(When clinician has finished remedial checklist, take both the diagnostic and remedial checklists and file. Give clinician his/her written diagnosis. Experimentor keeps written remediation.)

To complete Part Two, I will read your number remediation elements to you one at a time. For each one, I will ask you to look over your diagnosis and tell me the number and code of the diagnostic element.

Clinicians like yourself almost never make one-to-one association between diagnosis and remediation, that you, too, will have diagnostic elements which will not be associated with remedial elements.

I will now read your remedial element Number One to you. Please look through your diagnosis and tell me which, if any, of your diagnostic elements you associated with remedial element Number One. If you do give me a diagnostic number, or numbers, you are free to use that number, or numbers, again for other remedial elements.

TURN ON TAPE RECORDER

SESSIONS ONE AND TWO END HERE

SESSION THREE ONLY: BREAK

(Experimenter fills out Diagnostic and Remedial Element Form. Transfer numbers to: Remedial-Diagnostic Association sheet; Remediation Only sheet; and Diagnostic Only sheet.)

We will now begin the third and final part of this last session. Three times, now, you have given me the opportunity to observe you at work. I would now like to ask you to share with me some of the thought processes which you brought to bear on the diagnosis and remediation of this third case.

Here are your written diagnosis and remediation. (Give diagnosis/remediation.)

Here is a form which you will use as a guide in reflecting on your work. (Hand clinician Remediation-Diagnosis Association questions and alternative answers. Experimenter keeps Remediation-Diagnosis Association sheet. Columns have already been filled in from Diagnostic and Remedial Element form.)

You will recall that I read you your remedial elements and you told me which, if any, diagnostic elements you associated with them.

TURN ON TAPE RECORDER

You associated Remedial element number ____ with diagnostic element number _____. Look over these two elements to refresh your memory. (Give clinician time to look at both elements.) I'm going to ask you two questions about this association. The questions are listed on your guide sheet, together with a number of alternative answers. Please tell me which alternative, or alternatives, best captures your thinking about this association.

First, why did you choose remediation number ____? [Experimenter checks off alternative(s) read by the clinician.] Second, what is the association between the remediation and the diagnosis? [Experimenter checks off alternative(s) read by the clinician.] Do you have any comments to add? (Experimenter writes down any comments offered.)

The next association was remedial element ____ with diagnostic element _____. (Continue procedure used for first association until all associations have been accounted for.)

(If there are any remedial elements that are left unassociated go on to the next section: Remediation Only. If all Remediation elements have been associated with Diagnostic elements, go on to Diagnosis Only section.)

Remediation Only

Please look at your written remediation plan. (Hand clinician written remediation.) There were ____ remedial elements which you indicated were not associated with any diagnostic element or elements. I'm going to ask you one question about each. It is listed on this

sheet along with alternative answers which you will use as a guide in reflecting on your work.

(Give clinician Remediation Only guide sheet. Experimenter keeps Remediation Only form. Column has already been filled in from Diagnostic and Remedial Element Form.)

Please tell me which alternative, or alternatives, best captures your thinking about this remediation. Remedial element _____ is not associated with a diagnostic element because [experimenter checks off alternative(s) read by clinician.] Do you have any comments to add? (Experimenter writes down any comments offered.) Continue procedure until all unassociated remediations are accounted for.)

Diagnosis Only

Finally, please look at your written diagnosis. (Hand clinician written diagnosis.) There were _____ diagnostic elements which you indicated were not associated with any remedial element. I'm going to ask you one question about each. It is listed on this sheet along with alternative answers which you will use as a guide in reflecting on your work. (Give clinician Diagnosis Only guide sheet. Experimenter keeps Diagnosis Only sheet. Column has already been filled in from Diagnostic and Remedial Element Form.)

Please tell me which alternative, or alternatives, best captures your thinking about this diagnosis. Diagnostic element number _____ is not associated with a remedial element because _____. [Experimenter checks off alternative(s) read by clinician.] Do you have any comments to add? (Experimenter writes down any comments offered.)

(Continue procedure until all unassociated diagnoses have been accounted for.)

I'd like to give you a chance now to respond to several free-response questions. You can talk as much or as little as you like. I will audio tape your answers.

1. What guides you in gathering information about a child and in determining what should be in the diagnosis?
2. What do you see as being the relationship between diagnosis and initial treatment plan?
3. Do you have any comments about the sessions?

END OF SESSION THREE

APPENDIX C
CUE COLLECTION FORM

APPENDIX C

CUE COLLECTION FORM

Clinician: _____
Session: _____
Case: _____
Date: _____

Time In _____

Time Out _____

[illegible]

APPENDIX D
SAMPLE DIAGNOSIS

APPENDIX D
SAMPLE DIAGNOSIS

S = Strength; W = Weakness; Obs = Observation

1 Obs

Greg is the youngest of three children in his family. He has above

25

average intelligence. His major problem is his inability to quickly

3 W

identify basic sight words.

Practice

Greg's oral reading is poor. He correctly identifies the beginning of words. I'm not certain about his ability to use phonetic analysis independently.

APPENDIX E
DIAGNOSTIC CHECKLIST

APPENDIX E

DIAGNOSTIC CHECKLIST

Clinical Studies Project
Institute for Research on Teaching
Michigan State University

Clinician: _____
Session: _____
Case: _____
Date: _____
Time: _____
Location: _____

S = Strength; W = Weakness; O = Observation

	S	W	O		S	W	O
PHYSICAL				FAMILY			
Physical: General				Family: General			
Physical development				Home environment			
General health				Parent-school cooperation			
Vision: General				Parent educational background			
Vision: Acuity				Parent attitude toward school			
Hearing: General				Parent-child relationship			
Hearing: Acuity				Sibling relationship			
Speech articulation				Other			
Motor coordination							
Other:							
				PERCEPTION			
				Perception: General			
				Auditory memory			
				Auditory discrimination			
				Auditory sequencing			
AFFECTIVE				Auditory blending			
Affective: General				Visual memory			
Attitude toward school				Visual discrimination			
Attitude toward reading: Instructional				Visual sequencing			
Attitude toward reading: Independent				Visual motor integration			
Relationships with peers				Other			
Motivation for reading							
Behavior in the classroom							
Emotional adjustment				LANGUAGE			
Self-concept				Language: General			
Variety of interests				Listening vocabulary			
Ability to deal with new situations				Speaking vocabulary			
Other				Verbal fluency			
				Other			

S = Strength; W = Weakness; O = Observation

	S	W	O		S	W	O
SCHOOL				ORAL READING			
School: General				Oral reading: General			
Grade level placement				Oral reading: Rate			
Ability to work in a tutoring situation				Oral reading: Fluency			
Ability to work in a small group				Oral reading: Punctuation			
Ability to apply reading skills				Oral reading: Self-correction			
Attending behavior				Oral reading: Phrasing			
Intellectual potential: General				Oral reading: Intonation			
Intellectual potential: Verbal				Substitutions contextually acceptable			
Intellectual potential: Nonverbal				Omissions contextually acceptable			
Ability to do grade level work				Insertions contextually Acceptable			
Ability to read at grade level				Use of context to determine word pronunciation			
Appropriateness of instructional materials				Independent reading level: Fluency			
Other				Instructional reading level: Fluency			
				Other			
STRUCTURAL ANALYSIS				SILENT READING			
Word Analysis: General				Silent Reading: General			
Structural Analysis: General				Silent Reading: Rate			
Use of root words				Silent Reading: Comprehension			
Use of suffixes				Other			
Use of prefixes							
Use of syllables							
Blending word parts into whole words							
Other							

S = Strength; W = Weakness; O = Observation

[illegible]

APPENDIX F
SAMPLE REMEDIATION

APPENDIX F
SAMPLE REMEDIATION

S = Strength; W = Weakness; Obs = Observation; T = Treatment

1 T
Highlight the middle of similar words.

2 Obs
Hopefully he will be able to

3 W
discriminate among similar words. Tim is relying too heavily on the

4 S
beginnings of words. His motivation, though, is excellent.

Practice

Time needs help in identifying isolated sight words quickly. It would
be helpful to know what his interests are. He should use sight words
in his own stories. His visual acuity is normal.

APPENDIX G
REMEDIAL CHECKLIST

APPENDIX G
REMEDIAL CHECKLIST

Clinical Studies Project
Institute for Research on Teaching
Michigan State University

Clinician: _____
Session: _____
Case: _____
Date: _____
Time: _____
Location: _____

S = Strength; W = Weakness; O = Observation; T = Treatment

	S	W	O	T	Any Additional Remedial Comments?
VISUAL					
VISUAL: General					
DISCRIMINATION: General					
Discriminate visually similar words					
Differentiate similar words in context					
Attend to word endings					
Other					
MEMORY: General					
Work on sequential visual memory of letters, words, phrases, sentences					
Other					
VISUAL-MOTOR: General					
Gain visual-motor control (e.g., use templates)					
Other					
ACUITY: General					
Other					
AUDITORY					
AUDITORY: General					
DISCRIMINATION: General					
Discriminate specific speech sounds					
Discriminate pairs of similar sounds (e.g., p, b)					
Discriminate like pairs of words					
Produce specific speech sounds					
Other					

S = Strength; W = Weakness; O = Observation; T = Treatment

	S	W	O	T	Any Additional Remedial Comments?
AUDITORY (Continued)					
MEMORY: General					
Work on auditory memory of sounds, words, phrases, sentences					
Other					
ACUITY: General					
Other					
PHONETIC ANALYSIS					
PHONETIC ANALYSIS: General					
SYMBOL-SOUND RELATIONSHIPS: General					
CONSONANTS: General					
Drill on grapheme/phoneme correspondences of consonants/ consonant blends					
Identify consonant blends/ digraphs using regular word patterns (e.g., ship, chip, drip)					
Substitute initial consonant sound(s)					
Substitute final consonant sound(s)					
Other					
VOWELS: General					
Name the vowels					
Drill on grapheme/phoneme correspondences of vowel digraphs/ single vowels					
Drill on short/long vowel generalizations (e.g., mad-made)					
BLENDING: General					
Blend sound units into words					
Other					

S = Strength; W = Weakness; O = Observation; T = Treatment

	S	W	O	T	Any Additional Remedial Comments?
PHONETIC ANALYSIS (Continued)					
APPLICATION: General					
Apply phonics to attack unknown word					
Other					
SIGHT WORDS					
SIGHT WORDS: General					
Identify high-frequency sight words (context)					
Identify high-frequency sight words (isolation)					
Build sight vocabulary					
Use sight words in phrases					
Use sight words in language experience stories					
Other					
STRUCTURAL ANALYSIS					
STRUCTURAL ANALYSIS: General					
SYLLABLES: General					
Teach syllabication generalizations (e.g., vc/cv)					
Identify common syllables (e.g., ing, un, re, ly)					
Count number of syllables seen and heard					
Recognize common word parts (e.g., ack, ick, ous, tion)					
Blend syllables into whole word					
Other					
COMPOUND WORDS: General					
Form compound words from familiar words					
Other					

S = Strength; W = Weakness; O = Observation; T = Treatment

	S	W	O	T	Any Additional Remedial Comments?
STRUCTURAL ANALYSIS (Continued)					
ROOT WORDS: General					
Identify root words					
Other					
AFFIXES: General					
Recognize common prefixes/suffixes					
Add prefixes/suffixes to root words					
Learn prefixes/suffixes as sight words					
Other					
ORAL READING					
ORAL READING: General					
RATE: General					
Improve rate					
Other					
FLUENCY: General					
Improve fluency					
Improve phrasing					
Other					
COMPREHENSION					
COMPREHENSION: General					
INFORMATION GATHERING: General					
Demonstrate factual recall					
Remember supporting details					
Remember sequence of events					
Other					

S = Strength; W = Weakness; O = Observation; T = Treatment

	S	W	O	T	Any Additional Remedial Comments?
COMPREHENSION (Continued)					
MANIPULATIVE THINKING: General					
Make appropriate inferences					
Make appropriate generalizations					
Identify main idea(s)					
Other					
CONTEXT: General					
Use context (semantic/syntactic) as an aid to comprehension					
Other					
STUDY SKILLS: General					
Set purpose					
Adjust rate to suit purpose					
Adjust rate to suit content					
Other					
READING PLACEMENT					
READING PLACEMENT: General					
Place in independent level materials					
Place at best instructional level					
Other					
SPELLING					
SPELLING: General					
Incorporate spelling into word recognition activities					
Spell phonetic words by word chunks					
Spell non-phonetic words by relying on visual memory					
Spell sight words accurately					
Other					

S = Strength; W = Weakness; O = Observation; T = Treatment

	S	W	O	T	Any Additional Remedial Comments?
MOTIVATION					
MOTIVATION: General					
Involve student in setting own purpose					
Give student responsibility s/he can handle					
Provide high interest materials at student's level					
Provide feedback					
Provide praise					
Keep instructional periods short					
Provide charts to show growth					
Other					
CONSULTATIONS					
CONSULTATIONS: General					
Obtain eye, and/or ear, and/or speech evaluation					
Consult with school nurse					
Consult with parents					
Consult with school counsellor					
Consult with teacher					
Consult with school social worker					
Consult with school psychologist					
Consult with learning disabilities specialist					
Other					
OTHER					

APPENDIX H

REMEDIATION/DIAGNOSIS ELEMENT FORM: SESSIONS ONE AND TWO ONLY

APPENDIX H

REMEDICATION/DIAGNOSIS ELEMENT FORM: SESSIONS ONE AND TWO ONLY

Clinician: _____

Session: _____

Case: _____

Date: _____

Remediation & Diagnosis				Remediation & Diagnosis				Remediation & Diagnosis			
Rx#	Code	Dx#	Code	Rx#	Code	Dx#	Code	Rx#	Code	Dx#	Code

APPENDIX I

DIAGNOSTIC/REMEDIATION ELEMENT FORM: SESSION THREE ONLY

APPENDIX I

DIAGNOSTIC/REMEDIAL ELEMENT FORM: SESSION THREE ONLY

Clinician: _____
Session: _____
Case: _____
Date: _____

Remediation and Diagnosis

Debriefing

Comments

Rx#	Code	Dx#	Code	#1	#2
-----	------	-----	------	----	----

APPENDIX J

DIAGNOSIS ONLY/REMEDIATION ONLY ELEMENT FORM:

SESSIONS ONE AND TWO

APPENDIX J

DIAGNOSIS ONLY/REMEDIATION ONLY ELEMENT FORM:

SESSIONS ONE AND TWO

Clinician: _____

Session: _____

Case: _____

Date: _____

Remediation Only				Diagnosis Only			
Rx#	Code	Rx#	Code	Rx#	Code	Rx#	Code

APPENDIX K
REMEDIATION AND DIAGNOSIS GUIDE SHEET

APPENDIX K
REMEDICATION AND DIAGNOSIS GUIDE SHEET

1. Why did you choose this remediation?
 - a I am most familiar with it
 - b It is usually available
 - c I usually have success with it
 - d other
2. What is the association between the remediation and the diagnosis?
 - a The remediation is effective for this weakness
 - b The remediation is effective for this strength
 - c The remediation is effective for observations of this nature
 - d Other

APPENDIX L

REMEDIATION ONLY GUIDE SHEET AND RESPONSE FORM

APPENDIX L

REMEDICATION ONLY GUIDE SHEET AND RESPONSE FORM

REMEDICATION ONLY GUIDE SHEET

This remedial element is not associated with a diagnostic element because:

- a This remediation should always be included no matter what the specific diagnosis may be
- b The remediation is always appropriate for this type of student
- c I forgot to include the diagnostic element in the write-up
- d Other

APPENDIX M

DIAGNOSIS ONLY GUIDE SHEET AND RESPONSE FORM

APPENDIX M

DIAGNOSIS ONLY GUIDE SHEET AND RESPONSE FORM

Clinician: _____
 Session: _____
 Case: _____
 Date: _____

Dx#	This diagnostic element is not associated with a remediation because:	Comments
	☐ The diagnostic element cannot be remediated ☐ The diagnostic element does not require remediation ☐ I don't know a remediation for this diagnostic element ☐ I forgot to include the remediation in my write-up ☐ Other	

APPENDIX N
CALCULATION OF PHI CORRELATION

APPENDIX N
CALCULATION OF PHI CORRELATION

	+		-	
+	a (++)	b (+-)	a + b	
-	c (-+)	d (--)	c + d	
	a + c	b + d	N	

$$\text{Phi} = \frac{(a \times d - b \times c)}{(a + c) \times (b + d) \times (c + d) \times (a + b)}$$

The statistic is bounded by -1 (statements are in cells b and d only) and 1 (statements are in cells a and c only), only if the distributions in the marginals are equal. In all other cases, the maximum and minimum values will be less than 1 and greater than -1.

Phi maximum (>0) = maximum number of matches within cells a and c, holding marginals constant. Phi maximum (<0) = maximum number of matches within cells d and b, holding marginal constant.

Since the Phi coefficient is bounded by the marginal frequencies of the agreement table, correction is usually recommended.

$$\text{Phi corrected} = \frac{\text{observed value of Phi}}{\text{absolute value of Phi}}$$

Inter-Clinician Correlation

An example of a completed table is as follows:

Statements of Clinician A, SIMCASE Q	Statements of Clinician B, SIMCASE Q	Domain Statements (Diagnosis, Remediation or Cue)
S1	S1	S1
S2	S2	S2
S3	S7	S3
		S4
		S5
		S6
		S7

		Clinician A SIMCASE Q		
		+	-	
Clinician B SIMCASE Q	+	2 (a)	1 (b)	3
	-	1 (c)	3 (d)	4
		3	4	7

Intra-Clinician Correlation

An example of a completed table is as follows:

Statements of Clinician A, SIMCASE Q	Statements of Clinician B, SIMCASE Q	Domain Statements (Diagnosis, Remediation or Cue)
S1	S1	S1
S2	S2	S2
S3	S7	S3
		S4
		S5
		S6
		S7

Clinician i
SIMCASE Q
Form One

		+	-	
Clinician i SIMCASE Q, Form Two	+	2 (a)	1 (b)	3
	-	1 (c)	3 (d)	4
		3	4	7

APPENDIX 0
MASTER DIAGNOSIS AND REMEDIATION

APPENDIX O

MASTER DIAGNOSIS AND REMEDIATION

Case 3

1. His sight vocabulary and word recognition behaviors appear reasonably intact for his present grade placement. The slightly depressed scores (6.8+) are probably caused by:
 - 1) Insufficient independent reading. He is not a reader in the sense that he does very much of it. This alone might account for the slightly reduced sight vocabulary.
 - 2) Visual screening suggests that he is to some degree farsighted. This acuity problem might in turn account for the lack of practice described in 1 above.

REMEDIATION: He should be referred to an optometrist/ophthalmologist for complete visual testing. Also, parents and school should create a support system (library card, reading time and place, feedback mechanism such as parent discussion, etc.). That would encourage him to develop a reading habit. Both quantity and type of reading should be carefully followed for purpose of feedback until such time as he is able to sustain the habit on his own.

2. He has a significant problem with the higher level decoding skills and their application during reading. This problem is caused by:
 - 1) Some real confusion with multi-syllable - affixed words. He is not secure with many vowel phonograms used (as syllables) and does not give reasonable approximations to these sound values.

REMEDIATION: These responses can be taught through daily analysis of the multi-syllable words using worksheets - tape recorders - or teacher cues in the sound values. Extended reading behavior (see #1 sight voc.) should also aid in learning to sound these confused letter patterns.

- 2) An even greater problem is his inability to modify or adjust a sounded word into its spoken representation. Even when he correctly segments and sounds a multi-syllable word, he will often misplace a primary accent or use a long vowel sound value the actual word demands a short or r controlled. His final recognition seldom is adjusted in these areas leaving the word unrecognized or a nonsense form. He does not seem to systematically demand that sounded words be read in terms of his own meaning - speaking - listening vocabulary.

3. He has a serious problem with fluent texting of contextual material. This is caused by:

- 1) Inappropriate responses to punctuation.

REMEDIATION: Model the use and purpose of punctuation in reading. Use simultaneous sight reading of appropriate material (5-6 grade equivalent) in which he attempts to match the rhythm, cadence, and intonation of the model. Five minutes a day of such practice should be sufficient.

- 2) A suspected deficit in language development. He does not speak well in that his flow of language is not smooth. This appears during initial language learning and probably has some impact on his reading fluency.

REMEDIATION: Encourage verbal discussion using play reading in which he can rehearse smooth and fluent language behavior.

4. He has a reading comprehension problem specifically related to the meaning demands of sophisticated content related material. This is caused by:

- 1) His apparent inability to schematize or structure information for potential inferences and memory. He appears to process details (information) but does not easily frame these for inferred relationships or as a memory cue.

REMEDIATION: Preface reading assignments with self generated questions and/or insights that can serve as a guide and referent for the information and thinking that should take place during reading. This process might also be modeled with the teacher demonstrating through "talking out" while reading the referent frames that are generated. Careful questioning with referral to material for answer cues can also create the appropriate thought processes. This again is a demonstrated technique.

APPENDIX P

PROPORTIONAL AGREEMENT FIGURES

- (1) Diagnostic Statements
- (2) Remediation Statements
- (3) Cue Forms
- (4) Remediated Diagnoses
- (5) Remedial/Diagnostic Associations

APPENDIX P
PROPORTIONAL AGREEMENT FIGURES

The horizontal axis in all five figures represents, as proportions, the number of sessions at which types of statements were made. That is:

0.17 = made during one session

0.33 = made during two sessions

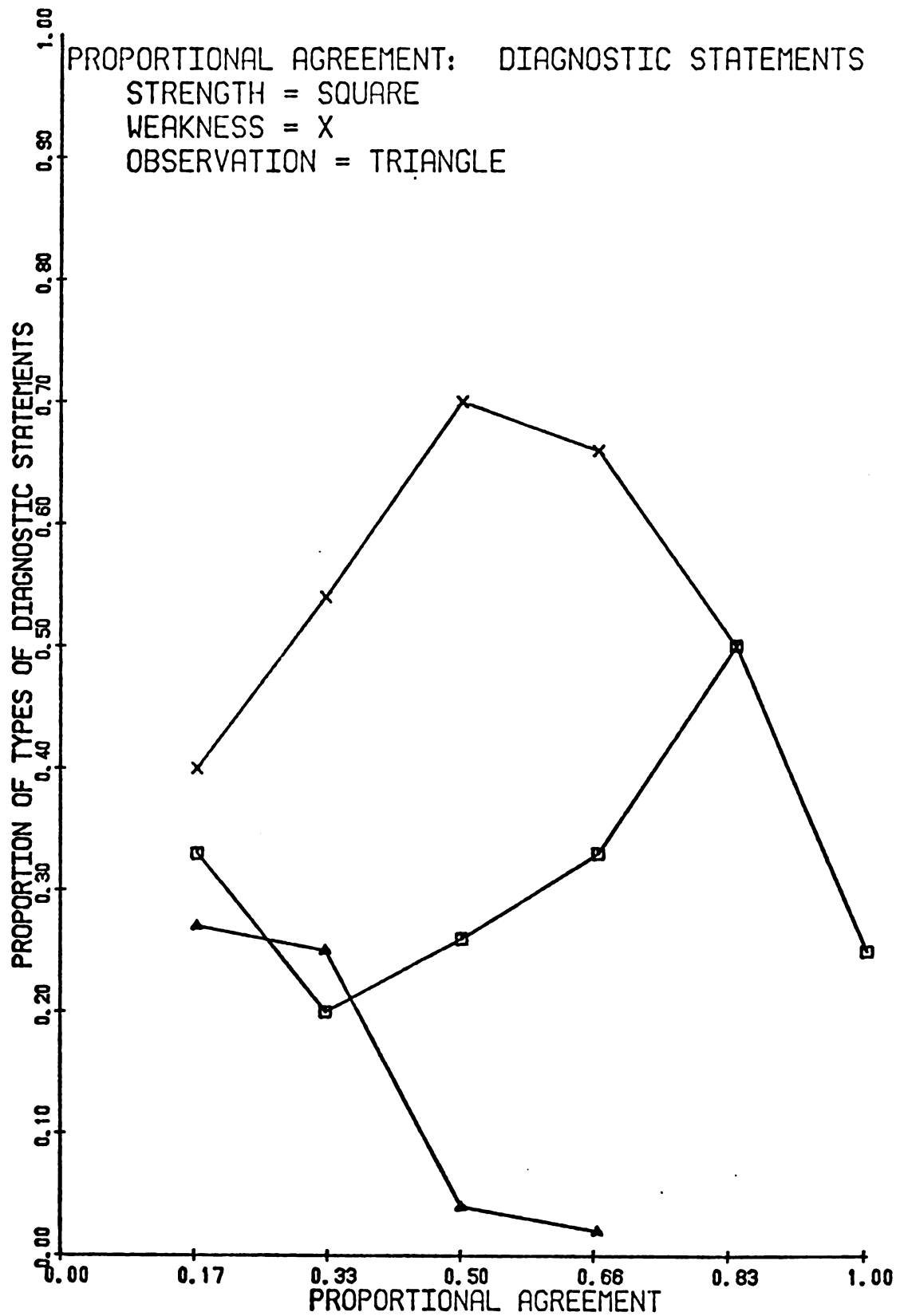
0.50 = made during three sessions

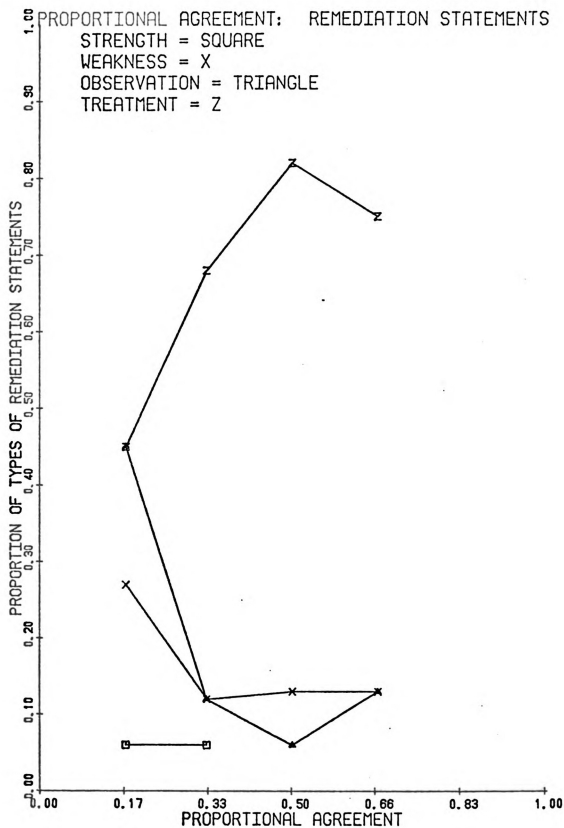
0.66 = made during four sessions

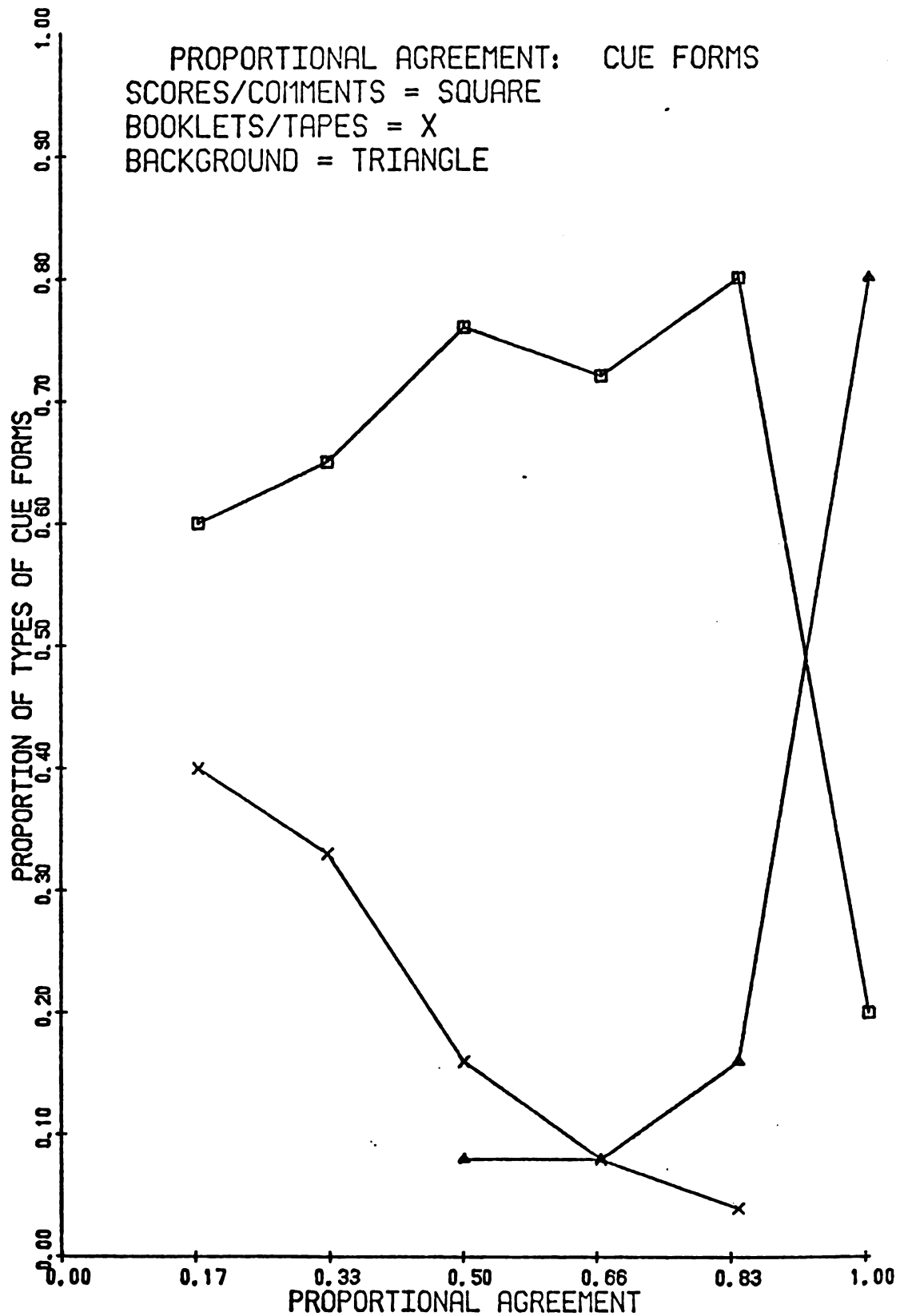
0.83 = made during five sessions

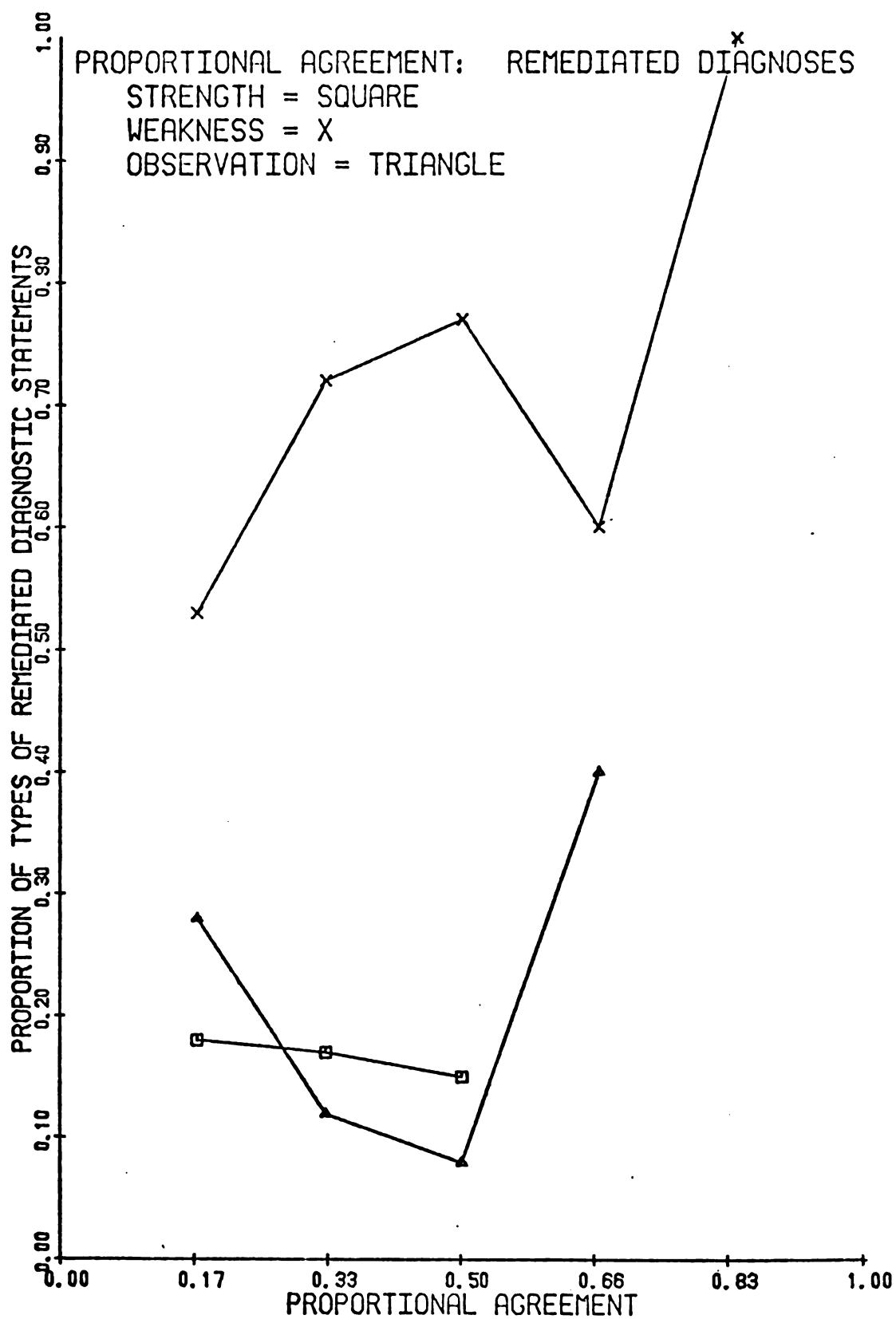
1.00 = made during all six sessions

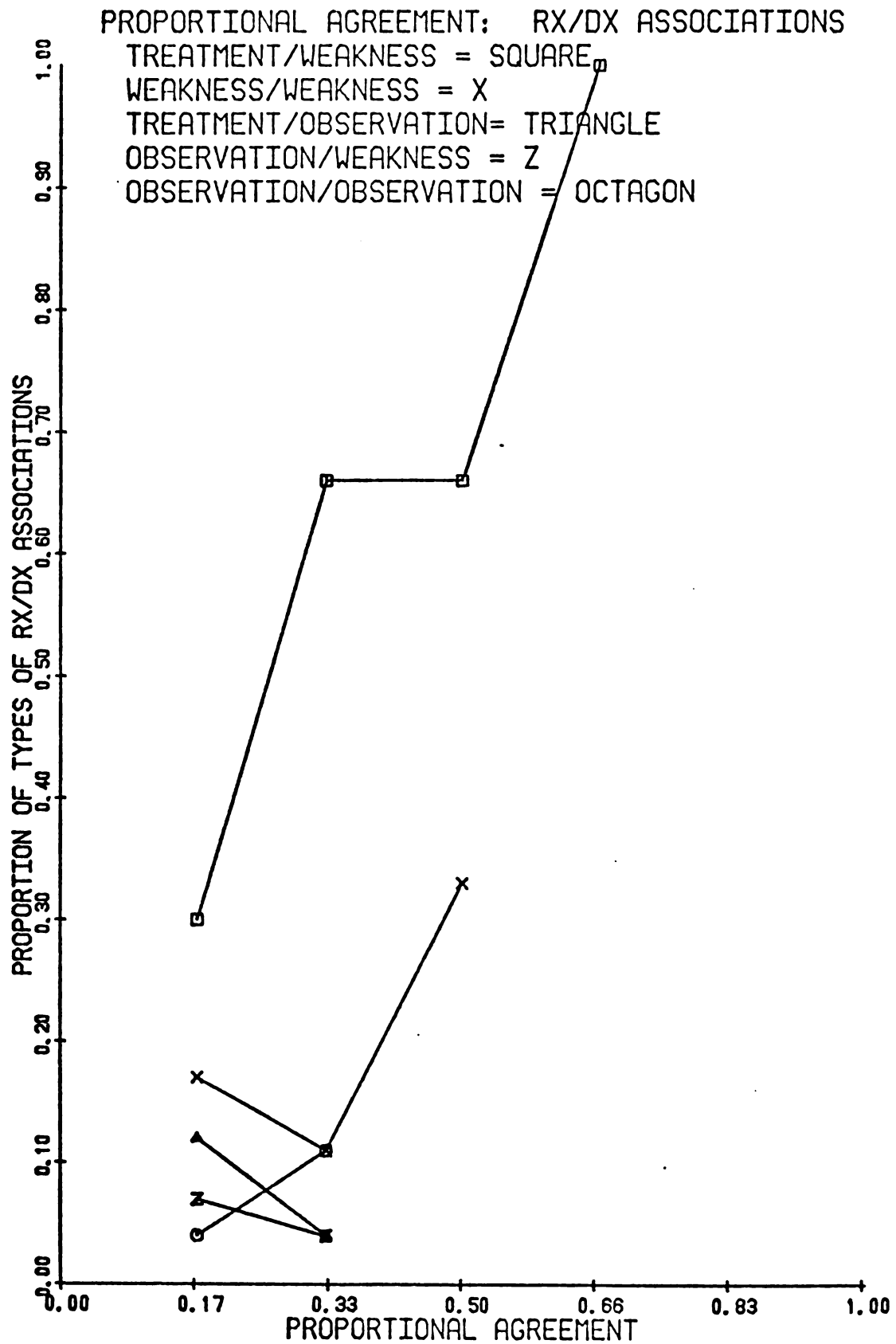
The vertical axis represents the proportion of each of the various types of statements made during each session. Looking at the figure for diagnostic statements, for example, statements coded as observations accounted for 0.25 (25 percent) of the diagnostic statements made in only one session (0.17 on the horizontal axis). However, statements coded as observations accounted for only 0.02 (two percent) of the diagnostic statements which were repeated across four sessions (0.66 on the horizontal axis).











APPENDIX Q
INTERCORRELATION RANGES

APPENDIX Q
INTERCORRELATION RANGES

(Porter Statistic is Shown in Parenthesis)

	Diagnosis		Remediation		Cue		Remediated Diagnoses		Remediated Associations	
	Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.
Case 1	0.05 (0.07)	0.43 (0.28)	0.03 (0.04)	0.43 (0.24)	-0.06 (0.15)	0.77 (0.77)	-0.03 (0.00)	0.35 (0.22)	-0.40 (0.00)	0.10 (0.08)
Case 2	-0.04 (0.0)	0.32 (0.21)	0.02 (0.03)	0.39 (0.29)	-0.04 (0.11)	0.68 (0.66)	-0.02 (0.00)	0.37 (0.18)	-0.35 (0.00)	-0.05 (0.03)
Case 3	-0.05 (0.0)	0.28 (0.20)	-0.04 (0.0)	0.29 (0.15)	-0.05 (0.15)	0.63 (0.56)	-0.03 (0.0)	0.21 (0.13)	-0.35 (0.00)	-0.10 (0.03)
Case 4	0.05 (0.05)	0.36 (0.24)	0.03 (0.04)	0.35 (0.24)	-0.06 (0.07)	0.61 (0.41)	0.04 (0.04)	0.33 (0.22)	-0.30 (0.00)	0.00 (0.02)

BIBLIOGRAPHY

BIBLIOGRAPHY

- Bateman, B. The role of individual diagnosis in remedial planning for reading disorders. In Calkins (Ed.), Reading Forum. NINDS Monograph No. 11. Bethesda, MD. 1971.
- Bond, G.L. Employ more diagnostic teaching in the classroom. Reading Improvement, 13(1), 35-39. Spring, 1976.
- Bond, G.L., & Tinker, M.A. Reading difficulties, their diagnosis and correction. New York: Appleton-Century-Crofts, 1967.
- Carter, H.L., & McGinnis, D.J. Diagnosis and treatment of the disabled reader. Toronto: MacMillan Co., 1970. (Chapters 3, 9.)
- Chall, J. Learning to read, the great debate. McGraw Hill, 1967.
- Chall, J., & Mirsky, A. Education and the Brain. In Chall & Mirsky (Eds.), The seventy-seventh Yearbook of the National Society for the Study of Education. Chicago: University of Chicago Press, 1978. (Chapter 11.)
- CEMREL, Inc. Reporting on Reading, Washington, D.C.: National Right to Read Effort, September 1979.
- Dechant, E. Diagnosis and remediation of reading disability. West Nyack, New York: Parker Publishing Co., June, 1969.
- Downing, J. Comparative reading. New York: MacMillan Co., 1973. (Chapters 4, 6.)
- Ekwall, E. Diagnosis and remediation of the disabled reader. Boston, MA: Allyn & Bacon, 1976.
- Elgin, S.H. The real literacy crisis. Change. November, 1978, 10(10).
- Elstein, A., Shulman, S., & Sprafka, S. Medical problem solving: An analysis of clinical reasoning. Cambridge: Harvard University Press, 1978.
- Flesch, R. Why Johnny can't read. Harper & Brothers, 1955.

- Gil, D., Hoffmeyer, L., VanRoekel, J., & Weinshank, A. Clinical problem solving in reading: Theory and research. (Res. Ser. No. 45). East Lansing: Institute for Research on Teaching, Michigan State University, 1979.
- Gil, D. The clinical problem solving behavior of classroom teachers as they diagnose a child's reading performance in laboratory and classroom. Unpublished doctoral dissertation, Michigan State University, 1979.
- Goodman, K.S. In Smith, F. (Ed.), Psycholinguistics and reading. Holt, Rinehart and Winston, 1973. (Chapter 14).
- Guthrie, J.T., & Siefert, M. In Myklebust, (Ed.), Progress in reading disabilities. Vol. IV.
- Guszak, F.J. Diagnostic reading instruction in the elementary school. Harper and Row, 1978. (Chapter 1.)
- Hall, M. & Ramig, C.J. Linguistic foundations for reading. Charles E. Merrill, 1978. (Chapter 7.)
- Harris, A.J. Five decades of remedial reading. Seattle, WA.: International Reading Association Annual Conference, May 5, 1967. (ERIC Document Reproduction Service No. ED 911830.)
- Harris, A.J. Ten years of progress in remedial reading. Journal of Reading. October, 1977, 21 (1), (29-35).
- Harvey, M. A good teacher and an eclectic approach: The hopeful answer to successful reading instruction. Reading Horizons. February, 1974, (15), (7-11).
- Hoffmeyer, L. The process and outcomes of diagnostic problem solving among reading clinicians. Unpublished doctoral dissertation, Michigan State University, 1980.
- Holmes, D.L. In Smith, F. (Ed.), Psycholinguistics and reading. Holt, Rinehart and Winston, 1973. (Chapter 5).
- Kendall, R.E. The role of diagnosis in psychiatry. Oxford, Great Britain: Blackwell Scientific Pub., 1975.
- Neurath, O., Carnap, R., and Morris, C. (Eds.). Foundations of the unity of science. (Vol. 2). Chicago: University of Chicago Press, 1979. (55-272).
- Lee, A., and Weinshank, A. Case production and analysis: CLIPIR pilot observational study of reading diagnosticians. (Res. Ser. No. 14). East Lansing: Institute for Research on Teaching, Michigan State University, 1978.

- Marcus, A. Diagnosis and accountability. Elementary English. May, 1974, 51 (5), (731-35).
- Melargno, R. Beyond decoding: Systematic schoolwide tutoring in reading. Reading Teacher. November, 1974, 28, (157-60).
- Natchez, G. (Ed.), Children with reading problems. Basic Books, Inc., 1968.
- Newell, A., & Simon, H. Human Problem Solving. Englewood Cliffs, New Jersey: Prentice Hall, Inc., 1972.
- Pearson, D. In Samuels, (Ed.), What research has to say about reading instruction. International Reading Association. 1978. (84-97).
- Robinson, H.M. Why pupils fail in reading. Chicago: University of Chicago Press, 1946.
- Rosenshine, B. & Mattleman, M. Effective remediation for low achievers in reading. Philadelphia, PA: Temple Univeristy, June, 1970. (ERIC Document & Reproduction Service No. ED045305).
- Rutherford, W.L. From diagnosis to treatment of reading disabilities. Academic Therapy. Fall, 1972, VIII (1), (51-55).
- Satz, P. In Otto, Peters, and Peters, (Eds.), Reading problems: A multi-disciplinary perspective. Addison Wesley, 1976. (Chapter 3).
- Smith, F. (Ed.), Psycholinguistics and reading. Holt, Rinehart and Winston, 1973.
- Smith, F. Understanding reading: A psycholinguistic analysis of reading and learning to read. Holt, Rinehart and Winston, 1978.
- Smith, N.B. American reading instruction. Newark, Delaware: International Reading Association, 1965.
- Smith, N.B. In Chall, J. (Ed.), Learning to read, the great debate. McGraw Hill, 1967.
- Spache, G.D. In Newman, (Ed.), Reading disabilities. Odyssey Press, 1969. (Chapter 2).
- Stauffer, R.G., Abrams, J.C., and Pikulski, J.J. Diagnosis, correction and prevention of reading disabilities. Harper and Row, 1978. (Chapter 2).
- Stein, J. (Ed.), The Random House dictionary of the English Language, the unabridged edition. Random House. New York, 1973.

- Strang, R. Reading diagnosis and remediation. Newark, Delaware: International Reading Association Research Fund, 1968. (Chapter 6).
- Stratoudakis, J. The construction and use of paper problems to observe the diagnostic problem solving behavior of reading clinicians. Unpublished doctoral dissertation, Michigan State University, 1980.
- Uhl, W.L. The use of the results of reading tests as bases for planning remedial work. Elementary School Journal. December, 1916, (266-275). XVII, (4), 266-275.
- VanRoekel, J. The problem solving behavior of reading specialists and learning disabilities resource teachers. Doctoral dissertation in progress, Michigan State University.
- Vinsonhaler, J. The consistency of reading diagnosis. (Res. Ser. No. 28). East Lansing: Institute for Research on Teaching, Michigan State University, 1979.
- Vinsonhaler, J., Wagner, C., and Elstein, A. The inquiry theory: An information-processing approach to clinical problem-solving research and application. (Res. Ser. No. 1). East Lansing: Institute for Research on Teaching, Michigan State University, 1978.
- Weinshank, A.B. The relationship between diagnosis and remediation in reading: A pilot study. (Res. Ser. No. 37). East Lansing: Institute for Research on Teaching, Michigan State University, 1979.
- Weinshank, A.B. The consistency of reading diagnosis. Paper presented at the annual meeting of the American Educational Research Association. Toronto, Canada. April, 1978.
- Weinshank, A.B. How reliable are clinical decisions in reading? Paper presented at the meeting of the Michigan Testing Conference. Ann Arbor, Michigan, March, 1980.
- Whipple, G.M. (Ed.), Report of the national committee on reading. Twenty-fourth Yearbook of the National Society for the Study of Education. Bloomington, Illinois: Public School Publishing Co., 1925. (Chapter 10).

MICHIGAN STATE UNIVERSITY LIBRARIES



3 1293 01754 3038