LEARNING FROM EXPERIENCE: APPLYING THE CLINICAL AND EPIDEMIOLOGICAL RESEARCH PARADIGMS TO THE STUDY OF DIAGNOSIS AND TREATMENT IN READING

By

Christian Carl Wagner

A DISSERTATION

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

DOCTOR OF PHILOSOPHY

Department of Counseling, Educational Psychology, and Special Education

1982

[illegible]

[illegible] wagne[illegible]

[illegible]

ABSTRACT

LEARNING FROM EXPERIENCE: APPLYING THE CLINICAL AND EPIDEMIOLOGICAL RESEARCH PARADIGMS TO THE STUDY OF DIAGNOSIS AND TREATMENT IN READING

By

Christian Carl Wagner

This research study examined the problem of how clinicians in any field learn from their experience. The traditional method for learning is through the informal aggregation of experience with its inherent problems of poor, biased human memory and inadequate variety. Alternate, formal methods, embodied in the clinical and epidemiological research paradigms, are universally accepted in the medical field as the basis for the continued improvement of medical care. These paradigms were, here, generalized to the diagnosis and treatment of problem readers. This study in reading was not intended to judge the ultimate adequacy of these paradigms for reading. This was made impossible by the limited variety of reading problems, diagnoses, and treatments examined. In the study all problems were quite severe, one cause was assumed and one treatment applied. Rather, the study was intended (1) to explicate the requirements for applying the research paradigms in reading, (2) to suggest refinements for future use of these methods, and (3) to document empirical relationships among children with reading problems, their

diagnoses, treatments.

The study was

classroom of one read

children. Reading pr

reeing the student pe

considered a major ac

ing, and post-treatme

in-class observation.

computer-assisted dia

data, diagnosis

epidemiological resear

treatment are the begin

result. The data was

cross-sectional time

perspectives. The major

examples of comput

valid diagnoses co

prevalence rates

seventh grade rea

descriptions of s

effectiveness ac

base-line assoc

diagnostic, trea

guidelines and s

diagnoses, treatments, and outcomes.

The study was conducted for one school year in the seventh grade classroom of one reading specialist who examined 42 and treated 30 children. Reading problems were of such a serious nature that even keeping the student performance from deteriorating further was to be considered a major accomplishment. Performance data was collected pre-, mid-, and post-treatment; the treatment itself was intensively coded by in-class observation. The data collected was used to generate reliable computer-assisted diagnoses and careful descriptions of treatment. Careful diagnosis was important since in the clinical and epidemiological research paradigms reliable diagnosis and differential treatment are the bedrock upon which valid learning from experience must be built. The data was analyzed in the prospective, retrospective, and cross-sectional time frames from both the clinical and epidemiological perspectives. The major results of the study included:

1. examples of computerized decision rules with which reliable and valid diagnoses could be generated;

2. prevalence rates for diagnostic profiles among severely deficit seventh grade readers;

3. descriptions of specific treatment plans and their differential effectiveness across various problem profiles;

4. base-line associations among students' personal, environmental, diagnostic, treatment, and outcome characteristics; and

5. guidelines and suggestions for the continued investigation of the

application of the

in reading.

across all of these me

differentially effect

clinical and epide

suited as means for

application of the clinical and epidemiological research paradigms in reading.

Across all of these results, it appears that reliable diagnoses and differentially effective treatments can be found in reading, and that the clinical and epidemiological research techniques should be further examined as means for learning from clinical experience in reading.

I were a p

could create a song.

one of these, I rea.

acknowledge and a

inadequate.

My family - Debb

My chairman - Dr

The rest of my c

Clarence VonTerson

And my other

Herman, Ms. Ruth

Waters

Acknowledgements

If I were a poet, I would write a poem. If I were a composer, I would create a song. If I were an artist, I would paint a picture. Being none of these, I leave it to my friends and loved ones to know how much I acknowledge and appreciate their help - for words alone are inadequate.

My family - Debbie, Brian, Andy, my parents, and so many more

My chairman - Dr. John Vinsonhaler

The rest of my committee - Drs. Arthur Elstein, Lee Shulman, and Lawrence VonTersch

And my other colleagues - Dr. Annette Weinshank, Dr. George Sherman, Ms. Ruth Polin, Mr. Ray Smith, Ms. Linda Vavrus and Ms. Debie Salters

Chris Wagner

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

Problem Statement *

Informal Aggregation of Experience

Learning from experience is a human ability often taken for granted. An individual encounters a problem, perceives certain characteristics of the situation, makes a decision, takes an action and observes the result. Somehow this experience is encoded into memory and serves to guide future decisions and actions. The learning is informal and very natural.

Reading specialists learn from their experience in this way. They are constantly confronted with problem readers. The specialists must decide what information to collect on the reader's environment and reading ability, what diagnostic classifications to use, what treatments to recommend and what outcomes to observe. The specialists learn by trying to perceive patterns of problems, treatments, and outcomes that confirm existing practices or suggest alternate ones. Each experience with a good reader or a poor reader can serve to inform and improve the specialist's diagnostic and remedial abilities.**

----

*The work reported herein is sponsored by the Institute for Research on Teaching, College of Education, Michigan State University. The Institute for Research on Teaching is funded primarily by the Teaching Division of the National Institute of Education, United States Department of Education. The opinions expressed in this publication do not necessarily reflect the position, policy, or endorsement of the National Institute of Education. (Contract No. 400-76-0073)

** Although the research described here is concerned with the diagnosis and treatment of reading disabilities, the general concepts should apply equally as well to any clinical problem solving field.

Unfortunate

experience is o

ties of human m

of the particu

quite narrow an

sult is incomplet

available in exper

direct opposition

indicate.

Given these

experience, it is

clinical experie

will learn prima

sult, the wea

through experienc

diagnostic and

across clinic an

of the field as

Although th

developed bett

clinical proble

After more tha

the epidemiolo

integration of

the nature of

and learning

Unfortunately, this informal method for learning from clinical experience is problematic (Einhorn, 1980, Einhorn & Hogarth, 1978). It relies on human memory, which may be quite faulty and biased; it relies on the particular set of cases the diagnostician encounters, which may be quite narrow and non-representative. The result is learning that, at best, is incomplete and does not make use of all information potentially available in experience. Far worse, the learning may actually be in direct opposition to what full and systematic use of experience would indicate.

Given these problems with informal methods of learning from experience, it is surprising that no alternate methods for learning from clinical experience seem to be in use in reading. Reading clinicians still learn primarily from their memory of the cases they see. As a result, the wealth of information available to an individual clinician through experience is not being effectively used to guide future diagnostic and remedial decisions. Furthermore, clinical experiences across clinicians are not being aggregated to guide the prevailing views of the field as a whole.

Although the field of reading diagnosis and treatment has not yet developed better techniques for learning from experience, another clinical problem solving field, medicine, does offer an alternative. After more than a century of development, it offers the formal clinical and epidemiological research paradigms for the aggregation and integration of clinical experience. These paradigms take into account the nature of human decisions and thinking and yield more reliable and valid learning from experience than do informal techniques. The

significance of
that they were de
response to s
experience. There
as effective gu
through the aggreg

The fields o
respect to the de
content of the p
leading clinic an
different experi
health of a patie
the clinician s
information abou
diagnosis, to su
the physician an
as possible.

Since phys
leading clinic
nature of inform
the history of
techniques that
methods. Blood-
century bu
research condu
, p.3.

significance of these paradigms for the field of reading diagnosis is that they were developed for use in a similar decision making setting in response to similar problems stemming from informal learning from experience. Therefore, it may be argued that these paradigms might serve as effective guides for improving reading diagnosis and treatment through the aggregation of clinical experience.

The fields of medicine and reading are, indeed, quite similar with respect to the decision making settings of their clinicians. Surely, the content of the physician's knowledge is vastly different from that of a reading clinician. However, the types of decisions made by these two different experts are the same. Each clinician is concerned with the health of a patient or case. The case has some problem which draws upon the clinician's area of expertise. The clinician is expected to collect information about the case and the case's environment, to make a diagnosis, to suggest a treatment, and to monitor outcomes. The goal for the physician and the reading clinician is to keep the case as "healthy" as possible.

Since physicians encounter the same decision making tasks as reading clinicians, they have also been confronted with the problematic nature of informal methods for learning from experience. Literature on the history of medicine is filled with examples of improper medical techniques that were followed because of reliance on these informal methods. Blood-letting as a medical treatment existed from at least the 11th century but was discontinued as a result of findings from formal research conducted by Louis (Garrison, 1929; Lilienfeld & Lilienfeld, 1980, p.34). The treatment of choice for gunshot wounds was

...ualization wi...

...rect comparis...

...pected prevent...

...lization unt...

...his was unwise...

...ractice was con...

...nd integrate exp...

...ranged clinical...

...ressions gather...

As these e...

...icine, informa...

...roblematic whet...

...are possible or...

...correct the p...

...rmal techniques...

...ormal Aggregatio...

The formal clinic...

...the experience...

...ormal techniqu...

...aradigms attemp...

...nd individual e...

...ure medical...

...cording and com...

...ontrasting cases...

cauterization with boiling oil until Pare found a better treatment by direct comparison of outcomes (Benger, 1961). More recently, the accepted preventive measure for swine flu included a special flu vaccination until careful collection of empirical data indicated that this was unwise (Neustadt & Fineberg, 1978). In each case, the improper practice was continued as long as informal means were used to aggregate and integrate experience. Formal methods provided information that changed clinical practice, information that was directly contrary to the impressions gained from informal reviews of experience.

As these examples illustrate, even in a "scientific" field like medicine, informal aggregation of experience is problematic. It is problematic whether the physician is highly motivated to give the best care possible or not. It has been problematic no matter how correct or incorrect the prevailing model of the human body was. Clearly, other formal techniques had to be developed for learning from experience.

Formal Aggregation of Experience

The formal clinical and epidemiological research paradigms for learning from experience are the medical field's response to the problems with informal techniques. Both the clinical and epidemiological research paradigms attempt to objectively and scientifically use the data from past individual experiences in clinical problem solving to improve future medical care. The paradigms direct the careful, systematic recording and comparison of characteristics across many similar and contrasting cases. They do not rely on human memory and continually fous of the

experiences. The

and Roberts, 19

These two re

problem solving

differences across

who has person

environment with

of these charac

reliable and va

physician must

administer it to

The clinica

to use the data

improve medical

These various da

The distribution

determining the

practices. Even m

The comparison of

comparisons can

clinical problem

The classi

epidemiological r

restricted. Epidem

focus on the problem of nonrepresentative or narrow clinical experiences. They have contributed greatly to the quality of medical care (Roberts, 1977).

These two research paradigms are intimately tied into the clinical problem solving model of medical diagnosis and the individual differences across patients. Medical care is concerned with a patient who has personal characteristics (denoted P), who lives in an environment with characteristics (E). The physician collects information on these characteristics in order to reach a diagnosis (D) that is a reliable and valid description of the true medical problems. The physician must then use this diagnosis to choose a treatment (T), administer it to the patient and monitor the outcome (O).

The clinical and epidemiological research paradigms, then, attempt to use the data (P,E,D,T,O) from individual clinical experiences to improve medical care. Both paradigms assume that the distributions of these various data in the medical setting are related and not random. The distribution of these characteristics across all cases can assist in determining the overall effectiveness of clinical problem solving practices. Even more important for improving clinical problem solving is the comparison of these characteristics across various groups. Such comparisons can indicate the differential effectiveness of alternate clinical problem solving procedures.

The classic division that exists between clinical and epidemiological research is defined by the data used and the data predicted. Epidemiology is concerned with explanation, with defining different diagno

presence or absen

part of the

personal and env

... Epi

environmental

medical diagnosis

problems,

patients and

research can,

physicians as we

health officials

Clinical

tion and outco

the factors af

the second pa

that maps the c

treatment tha

research searc

patient's outco

cover the

data. Since th

the ou

clinical resea

medical prob

the choice of

different diagnostic categorizations and the factors affecting their presence or absence. For the practicing physician, this is the first part of the clinical problem solving process, the part that maps the personal and environmental characteristics of the case into a diagnosis (P,E -> D). Epidemiology searches for relationships among the personal and environmental characteristics of the patient (P,E) and the patient's medical diagnosis (D). Since one diagnostic possibility must always be "no problem," epidemiological research examines both patients and non-patients and focuses on the "causes" of medical problems. This type of research can, then, be used to assist in the diagnosis of patients by physicians as well as in attempts to prevent medical problems by public health officials.

Clinical research, on the other hand, is concerned primarily with action and outcome, with defining different treatment alternatives and the factors affecting their success. For the practicing physician, this is the second part of the clinical problem solving process, the part that maps the characteristics of the case and its medical diagnosis onto a treatment that produces the best outcome (D -> T -> O). Clinical research searches for relationships among the treatment (T) and the patient's outcomes (O) for a given diagnosis (D) in an attempt to discover the most effective treatment for a given profile of patient data. Since this research is concerned with a specific kind of outcome, i.e., the outcome of a treatment applied to a particular problem, clinical research examines only patients and focuses on the treatment of medical problems. This type of research, then, can be used to assist in the choice of treatments for particular patient profiles and in the statement of a

epidemiological

studies.

The protot

1854 identified

Snow had noted

London where water

Southwark and Va

from their water

from London. Sn

the transmission

Then in 1854,

conditions and o

Environmental
Characteristics
Supplier

Southwark & Vau

Lambeth Co.

Other

Note that the

houses serviced

that company

sections of the

house to anot

statement of a prognosis. The differences between clinical and epidemiological research can be clearly seen in the following two examples.

The prototypic epidemiological study was conducted by John Snow in 1854 (Lilienfeld & Lilienfeld, 1980). During a cholera epidemic in 1849, Snow had noted that cholera rates seemed particularly high in areas of London where water was supplied by two specific water companies, the Southwark and Vauxhall Company and the Lambeth Company. These suppliers drew their water from a very polluted section of the Thames downstream from London. Snow suspected that this contaminated water was implicated in the transmission of cholera, possibly through some intestinal germ. Then in 1854, another cholera epidemic began. Snow decided to test his suspicions and obtained the data summarized in Table 1.

Table 1
The Classic Epidemiological Study

Environmental Characteristic: Water Supplier	Number of Houses	Diagnosis: Deaths from Cholera	Deaths per 10,000
Southwark & Vauxhall Co.	40,046	1,263	315
Lambeth Co.	26,107	98	37
Other	256,423	1,422	59

Note that the death rate from cholera was 5 to 10 times higher in the houses serviced by the Southwark and Vauxhall Company than elsewhere. That company still drew its water from some of the most polluted sections of the Thames. The Lambeth company had since changed its water source to another part of the Thames upstream of London where the other

...ries were

...ween an env...

...egory the e...

...water became a...

of cholera sugg...

...ears, laws were

...ter the... wa...

...are of an ind...

...ther, with...

...resence or abse...

A typical

...leagues (Bu...

...firmed breas...

for 5 years.

...emotherapy w...

...acebo. The re...

...CMFVP adju...

...acebo

...e data ...

...emotherapy.

...atistically

companies were also situated. Snow had shown such a strong relationship between an environmental characteristic (water quality) and a diagnostic category (cholera) that the existence of a "cholera poison" transmitted by water became an accepted fact. This new understanding about the cause of cholera suggested a rather obvious, preventive measure. Within two years, laws were passed requiring all water companies in London to filter their water. Note that this research did not address the proper care of an individual who had contracted cholera but was concerned, rather, with understanding cholera and the factors affecting its presence or absence.

A typical clinical research study was conducted by Fisher and colleagues (Burdette & Gehan, 1970, p.39). Fifty-nine women with confirmed breast cancer treated with radical mastectomy were followed-up for 5 years. During this time 23 of the 59 received adjuvant chemotherapy with the drug thio-TEPA and the remaining 36 received a placebo. The resulting data are given in Table 2.

Table 2
A Typical Clinical Research Study

	Recurrence of cancer within 5 years	No recurrence of cancer within 5 years
Thio-TEPA adjuvant therapy	65% (15 of 23)	35% (8 of 23)
Placebo	86% (31 of 36)	14% (5 of 36)

The data clearly suggest the therapuetic possibilities of the chemotherapy. After analysis, the differences in recurrence rates were statistically significant. Note that this research addressed the problem of treating pat

related to the pr

in the first p

Generalizing

research to rea

clinical prob

strategies. The

strategies would

across individu

the empirical

reality. Readers

clinician decis

these decision

determine and

problem solving

Solving Forma

The successes

medicine and th

medicine sugge

in reading cou

"formal" metho

involved care

attempted to

strategies from

of treating patients with cancer rather than understanding the factors related to the presence or absence of such cancer and trying to prevent it in the first place.

Generalizing the medical paradigms of clinical and epidemiological research to reading diagnosis seems straightforward. The concepts of clinician, problematic case, diagnosis, treatment, etc., have direct analogues. The generalized clinical and epidemiological research paradigms would direct the aggregation of empirical data within and across individual clinicians and readers. The data would be used to relate empirically the particular characteristics of problematic and healthy readers, available diagnostic and treatment possibilties, clinician decision making, and changes in case performance following these decisions. These empirical associations could then be used to determine and improve the effectiveness and efficiency of clinical problem solving performance in reading.

Applying Formal Methods for Aggregating Experience to Reading

The successes of the clinical and epidemiological research paradigms in medicine and the apparent similarity between the fields of reading and medicine suggested that the generalization of these research paradigms to reading could serve as an effective alternative to the current informal methods of learning from experience; they could lead to improved care for problem readers. Therefore, the research reported here attempted to generalize the clinical and epidemiological research paradigms from medicine to the field of reading diagnosis. This was accomplished by

in reading. Th

clinical diagno

readers. The re

their performance

would be conside

were recorded so

characteristics

would be determ

cooccurrences

characteristics

means for iden

Since this

clinical and ed

the results were

paradigms for

is is true fo

medicine, the

the number and v

techniques. The

constrained by th

treatments exam

in which one

prescribed. It

the way that the

the clinical and

accomplished by conducting a clinical and epidemiological research study in reading. The study was run in a setting in which one reading clinician diagnosed, treated, and followed up a number of problem readers. The readers examined had such severe problems that preventing their performance from falling further behind their grade placement would be considered a major success. Many characteristics of the setting were recorded so that the relationships among personal and environmental characteristics (P,E), diagnoses (D), treatments (T), and outcomes (O) could be determined and compared across naturally occuring groups. Then, co-occurrences of case, clinician, treatment, and outcome characteristics were counted, compared across groups, and interpreted as a means for identifying the more reasonable decision alternatives.

Since this study was apparently the first attempt to generalize the clinical and epidemiological research paradigms in the field of reading, the results were not expected to rule on the ultimate adequacy of the paradigms for reading. Clearly, no single study could be so persuasive. As is true for clinical and epidemiological research studies in medicine, the outcomes of this study were attenuated by limitations on the number and variety of cases examined and by the chosen research techniques. The generalizability of the results of this study was restrained by the limited variety of reading problems, diagnoses, and treatments examined; the reading problems were all of a severe nature for which one common cause was assumed and one basic treatment prescribed. It was only after a century of medical studies limited in this way that the medical field came to universally accept the tenets of the clinical and epidemiological research paradigms as a foundation for the continued

rarely come to such

However, this

requirements of the

to meet these requ

research were the

examination of th

paradigms will be

very common sense

recording of events

This study,

to perform a clinic

problem solving

field of reading d

evaluate the res

research paradigms

reading diagnosi

understanding of

medicine. Specif

student characte

environmental cha

diagnoses, treatme

the use of the two

the continued improvement of medical care. The field of reading could hardly come to such acceptance on the basis of one study.

However, this initial study did provide many lessons on the precise requirements of the research paradigms and the reading field's ability to meet these requirements. Possibly the most important results of this research were the more explicit requirements for the continued examination of the research paradigms in reading. Whether or not the paradigms will be ultimately acceptable cannot be speculated, but the very common sense notions of learning from experience by systematic recording of events deserve further scrutiny.

This study, then, had two purposes. The first was to determine how to perform a clinical and epidemiological research study in any clinical problem solving field in general, and to carry out such a study in the field of reading diagnosis in particular. The second purpose was to evaluate the results of applying the clinical and epidemiological research paradigms in reading with respect to further understanding of reading diagnosis and treatment and with respect to further understanding of the task of generalizing the paradigms outside of medicine. Specific topics were: 1)baseline likelihooods of significant student characteristics, 2)associations between personal or environmental characteristics and diagnoses, 3)associations between diagnoses, treatments, and outcomes, and 4)problems and concerns with the use of the two research paradigms in reading.

The purpose of this s

clinical problem

informal aggregation

problem solving qua

clinical and epidem

experience, and

teaching with partic

process.

The Clinical Proble

The clinical prob

content areas (cust

with the operation

example, medical

patients; readin

electronic trouble

automobiles. In ea

presenting any pe

area. These area

problems, but the

similar.

CHAPTER 2

Related Research

The purpose of this chapter is to document (1) the characteristics of the clinical problem solving task, (2) the problems inherent in using informal aggregation of experience to determine and improve clinical problem solving quality, (3) the methods and past successes of the formal clinical and epidemiological research paradigms for aggregating clinical experience, and (4) the current state of clinical problem solving in reading with particular emphasis on models of the reader and reading process.

The Clinical Problem Solving Model

The clinical problem solving model is appropriate for many different content areas (Lusted, 1968, Preface). It is used in fields concerned with the operation, maintenance and improvement of complex systems. For example, medical clinicians work to solve the problems of their patients; reading specialists - the problems of their readers; electronic trouble shooters - their circuits; auto mechanics - their automobiles. In each case there is an expert responsible for solving and preventing any performance problems of systems in his/her respective area. These areas may differ vastly in the characteristics of their systems, but the thinking processes of the experts are remarkably similar.

The major cog

setting are clear.

complex system ca

with interact

toward achieving so

can be inferred f

set of Critical Pe

upon the internal

The second pa

clinician. The c

case's performance

set of actions

performance and u

parts interact.

The interact

a schedule for

case performance

participants are

The first th

characteristics

existing or poten

believed to a

includes the pr

evaluate impro

organized arou

The major cognitive characteristics of the clinical problem solving setting are clear. There are two participants. The first is some type of complex system called a "case." The case can be viewed as a set of parts which interact through some internal process and which are directed toward achieving some goal. It is assumed that the "health" of the case can be inferred from its ability to demonstrate competence on a certain set of Critical Performances. The case's perceived competence depends upon the internal process by which the various case parts interact.

The second participant is an expert clinical problem solver, a clinician. The clinician has the task of maintaining and improving the case's performance. The clinician's ability to do this depends upon the set of actions that can be initiated by the clinician to improve case performance and upon the clinician's own model of how the various case parts interact.

The interaction between clinician and case is usually initiated by a schedule for routine and preventive maintenance or by a problem with case performance (DeGowin & DeGowin, 1976, Chapters 1-2). Once the participants are together, clinical problem solving begins.

The first thing the clinician must do is collect information on the characteristics of the case and its environment in order to detect existing or potential problems. The clinician then selects a procedure believed to alleviate or prevent the problems. Finally, the clinician executes the procedure and monitors the case's response in order to evaluate improvement. These actions taken by the clinician have been organized around the terms "diagnosis," "treatment," and "followup." For clinician, the

problematic case

to help select the

the problematic

wife, 1972, Chapter

reading - Spache,

Treatment is the

clinician to improve

& Spache, 1976,

case performance

adjustment of dia

changes Small & M

is directed by the

clinician.

But the situa

way. At least

Kinsbourne, we

imply a statem

of Critical Pe

performance to th

use of diagn

performances are

The second

case's problematic

Critical Perfor

establish the sta

a clinician, the term "diagnosis" is defined as the statement of any problematic case performances and their causes. This statement is used to help select treatments and is based upon all information known about the problematic case (in medicine - Wulff, 1976, Chapter 7; Barnoon & Wolfe, 1972, Chapter 2; in epidemiology - Friedman, 1980, Chapter 13; in reading - Spache, 1976, Chapter 1; in system design - Hare, 1967). "Treatment" is the specification of a set of actions initiated by the clinician to improve or sustain case performance (Roberts, 1977, Chapter 4; Spache, 1976, Chapter 1). "Followup" is the monitoring of changes in case performance that occur during and after treatment and the adjustment of diagnostic and treatment decisions based upon these changes (Small & Krause, 1972). The clinical problem solving task, then, is directed by the diagnostic, treatment, and followup decisions of the clinician.

But the situation is more complex than these simple definitions imply. At least two different levels of diagnosis are often used (Vinsonhaler, Weinshank, Wagner, & Polin, 1982). The first level is simply a statement of problematic functioning. The clinician examines all Critical Performances* of the case and compares the case's actual performance to the case's expected or optimal performance. This first level of diagnosis is often no more than a statement of <u>what</u> performances are problematic.

The second level of diagnosis is a statement of the cause of the case's problematic Critical Performances. These causes may include other

----

* Critical Performances are a set of criterial evaluations used to establish the state of case performance.

concurrent charact

performance and ex

performances are p

This second

perform than the f

expert clinician to

disease process mus

the history of c

performances. In sc

very well elabora

system, electrica

each of these, and

more ably and va

ability, speed,

and treatment of p

valid models.

experience, home

instruction rece

related to Critica

oral reading flu

agree on a definit

Given these t

distinguished. One

systematic Criti

treatment. The sec

of the problem. Wh

concurrent characteristics of the case as well as the case's history of performance and experience. This second level of diagnosis describes <u>why</u> performances are problematic.

This second level of diagnosis is generally more difficult to perform than the first level. The second level diagnosis requires the expert clinician to have and use a model of the case process. This model of case process must relate the functioning of internal case parts and the history of case performance and experience to the case's Critical Performances. In some fields like automobile repair, the model can be very well elaborated and valid. The various internal parts (coolant system, electrical system, drivetrain, etc., and the subcomponents of each of these) and history (miles driven, mainenance provided, etc.) can be reliably and validly related to the Critical Performances (steering ability, speed, acceleration, etc.). Other fields, like the diagnosis and treatment of problem readers, may have very fuzzy, unreliable and invalid models. Various internal aspects of the case (language experience, home environment, attitude, etc.) and its history (instruction received, etc.) are only unreliably and probabilistically related to Critical Performances (silent reading rate and comprehension, oral reading fluency, etc.), assuming, of course, that the field can agree on a definition for the important Critical Performances!

Given these two levels of diagnoses, two types of treatments can be distinguished. One type of treatment would be aimed at alleviating the problematic Critical Performances directly - a so called symptomatic treatment. The second type of treatment would be directed at the cause of the problem. Whatever the type of diagnosis or treatment, the goal is to improve the

uring the admis

xpertise to com

rescribed. Si he m

is expected and

mprovements may le

roblem solving pro

deal with the case

ility to demonst

To summarize,

ternalized mode

ctions about the

1. determining t
P),

2. determining
case (E),

3. determining
performances

4. choosing the

5. monitoring a
determine ou

The basic cha

ntent area kno

to improve the case's problematic Critical Performances. To this end, during the administration of treatment, the clinician uses his/her expertise to compare the actual treatment received to the treatment prescribed. S/he must further compare the case's actual improvement to its expected and optimal improvement. Any large discrepancies in these improvements may lead to a repeat of the entire diagnostic and remedial problem solving procedure with one difference - the diagnosis would then deal with the case's ability to change with treatment as well as its ability to demonstrate its Critical Performances.

To summarize, the clinician is a problem solver. S/he has an internalized model of case functioning organized around decisions and actions about the following:

1. determining the significant personal characteristics of the case (P),

2. determining the significant environmental characteristics of the case (E),

3. determining the diagnosis of the case (D), both as to what Critical Performances are problematic and why they are problematic,

4. choosing the treatment for the case (T), and

5. monitoring actual case change and comparing it to expected change to determine outcome (O).

The basic chain of reasoning is P,E -> D -> T -> O. The clinician uses content area knowledge to make decisions concerning the measurement procedures to use

reject or accept

outcomes to expect

clinician's goal

functioning as opt

Informal Aggregatio

Determining an

during just descr

decisions can only

similar decisions

possible to define

outcome data can

experience. Theref

based upon data

competence must be

for improvement.

clinical experien

in practicing cl

Typically, a

particular popu

clinical decision

changes, and tr

cases and clinic

estimates of

procedures to use next; the diagnostic classifications to consider, reject, or accept as case descriptors; the treatments to administer; the outcomes to expect; etc. (Elstein, Shulman, & Sprafka, 1978). The clinician's goal is to keep the cases in his/her area of expertise functioning as optimally as possible within cost/benefit constraints.

Informal Aggregation of Experience

Determining and improving the quality of the clinical problem solving just described is a difficult task because the quality of these decisions can only be evaluated by comparison of outcomes across other similar decisions. Only when such relative outcomes are available is it possible to define better and worse clinical decisions. But relative outcome data can only be obtained through the aggregation of clinical experience. Therefore, any attempt to improve clinical decisions must be based upon data from clinical experience. Existing levels of clinical competence must be "learned" from clinical experience, as must methods for improvement. By far the most common method for gathering and using clinical experience is the informal aggregation performed naturally by any practicing clinician.

Typically, a clinician only diagnoses, treats, and follows-up the particular population of cases that make up his practice. Specific clinical decisions are only <u>informally</u> connected with case performance changes, and these connections are only informally aggregated across cases and clinicians. As a result there are no empirical, objective estimates of decision making reliability or validity, of case performance outco

outcomes. There ar

groups of clinic

diagnostic categor

results. Because

involvement of cl

difficult. Alternat

meaningfully compa

These proble

part to the struct

has limited time

Therefore, many c

training and exp

decisions. But, e

clinician will

representative, s

the clinician s

population of ca

decisions (Rober

But even pr

clinicians make

Einhorn, 1980,

the inappropria

of information

are not made.

decision that is

performance outcomes, or of the associations of decisions and case outcomes. There are no estimates for individual clinicians or across groups of clinicians. Treatments of unspecified effect are applied to diagnostic categories with unknown reliability to give uncertain results. Because of the fuzziness in relating decisions to outcomes, improvement of clinical problem solving performance is extremely difficult. Alternative decisions and their respective outcomes cannot be meaningfully compared.

These problems with informal aggregation of experience are due in part to the structure of typical clinical practice. A given clinician has limited time and can only interact with a finite number of cases. Therefore, many circumstances may exist for which the clinician's training and experience are not broad enough to make valid and reliable decisions. But, even when clinical experience is broad enough, a given clinician will only interact with the particular, and not necessarily representative, set of cases that comprise his/her practice. Therefore, the clinician's experience may not be representative of the true population of cases leading, again, to unreliable or invalid clinical decisions (Roberts, 1977).

But even broad and representative experience may prove inadequate. Clinicians make decisions that lead to actions and then observe results (Einhorn, 1980). So called "false positives," reaching decisions that are inappropriate, are observable and can be instructive. Unfortunately, no information is available for particular alternative decisions that were not made. The "false negative," not reaching a particular alternate decision that is appropriate, is not observable (Einhorn & Hogarth, [illegible]. Experience.

procedures that

alternatives. Two

than optimal learn

cases and decision

a less stringent an

Broad, repres

not a guarantee

because of the p

experience. E

decision makers

from experience

perspective. Then

cases that aris

of experience.

probability judg

clinician consid

the efficacy o

"successes" in m

procedure to b

often confirm th

[illegible], 1974).

particular succ

ability, that i

But there i

patterns with

1978). Experience, then, is generally limited to refining existing procedures that seem to work and does not promote the testing of new alternatives. Typical clinical practice can, therefore, lead to less than optimal learning from experience. The requirements for the types of cases and decisions that must be experienced for ideal learning are quite stringent and are rarely met informally.

Broad, representative experience on all possible decisions is still not a guarantee that informal learning from experience will be valid because of the biased ways in which clinicians might interpret their experience. Einhorn (1980) reports research indicating that human decision makers ignore false positives and false negatives in learning from experience - even when they are available. From another perspective, Tversky and Kahneman (1974) describe a large number of biases that arise in human thinking that could improperly alter the use of experience. For example, human problem solvers tend to make probability judgements on the availability of instances in memory. As a clinician considers a particular decision, s/he might well overestimate the efficacy of a procedure due to the high availability of his/her "successes" in memory. Phrased differently, if the clinician expects a procedure to be effective, s/he will tend to remember only those cases which confirm the expectation - the phenomenon of illusory correlation (Slovic, 1974). Furthermore, a good match between a particular situation and similar successfully completed situations can lead to an illusion of validity, that is, unwarranted confidence in results.

But there is still one more problem. Assuming that all of the concerns with informal aggregation of experience described above have been addressed, s

the clinician's de

problem solving

be very unreliable

clinician alters

experience, the cl

The results of

unreliable, can be

Assume that

described in Tab

Case Number

1
2
3
4
5
6
7
8

Because of the u

all cases as h

the clinician wa

cases supposed

the true problem

estimate of the

the effects of t

been addressed, such informal aggregation may still be a poor teacher if the clinician's decisions are unreliable. Research in many clinical problem solving fields has demonstrated that human decision making can be very unreliable (Yerushalmy, 1972; Vinsonhaler et al, 1982). As the clinician alters his/her procedures for decision making based on experience, the clinician assumes that his/her decisions are reliable. The results of assuming reliability when, in fact, the decisions are unreliable, can be seen in the following example.

Assume that an unreliable clinician has treated the set of cases described in Table 3.

Table 3
Results of Unreliable Diagnosis

Case Number	True Diagnosis	Clinician's Unreliable Diagnosis	Selected Treatment	Outcome
1	A	A	X	Outcome-1
2	B	A	X	Outcome-2
3	C	A	X	Outcome-3
4	A	A	X	Outcome-1
5	A	A	X	Outcome-1
6	B	A	X	Outcome-2
7	D	A	X	Outcome-4
8	A	A	X	Outcome-1

Because of the unreliable decisions, the clinician incorrectly diagnoses all cases as having problem A and uniformly prescribes treatment X. As the clinician watches the outcomes of applying treatment X to these cases supposedly with problem A, s/he does not take into account that the true problem of many of the cases was not A. The clinician's estimate of the effect of treatment X on problem A is inaccurate because the effects of treatment X on problems B, C, and D confound it. Clearly, the lower the re

frequent the w

estimating the use

As this e

information that c

experience may be

clinicians make de

The implicat

solving and human

aggregation of

determination or

relative efficienc

estimated. Alternat

reasonably compare

reaching diagnosis),

will be taught p

be as accurate as

than optimal clini

What is needed

empirical clinica

decision making re

limitations of h

non formal system

epidemiological re

techniques provide

the lower the reliability of decision making (and, therefore, the more frequent the misdiagnosis), the greater the potential for error in estimating the usefulness of treatment X for problem A.

As this example suggests, even the biased and incomplete information that can be had from the informal aggregation of clinical experience may be of limited validity given the unreliable way in which clinicians make decisions.

The implication of these characteristics of clinical problem solving and human thinking is that existing informal procedures for the aggregation of experience do not lend themselves to the valid determination or improvement of decision making quality. Neither relative efficiency nor relative effectiveness can be meaningfully estimated. Alternative decision making strategies cannot, therefore, be reasonably compared. When this is true in a field (e.g., medical or reading diagnosis), even training programs will be affected. Students will be taught prevailing views of good clinical practice that may not be as accurate as possible, starting the new clinicians out with less than optimal clinical problem solving expertise.

What is needed, then, is a formal system for the aggregation of empirical clinical experience, a formal system that would guarantee decision making reliability and eliminate, as much as possible, the limitations of human thinking and individual clinical experience. One such formal system is available through the techniques of clinical and epidemiological research. According to Roberts (1978, p. 4), these techniques provide "possibly, the only methodological framework within with science and

management of inc

Formal Aggregation

Research Paradigms

The clinical

problem of deter

aggregation of ex

considered sub-fi

techniques can, in

content area. Bot

collecting persona

making a diagnosis

observing outcomes

Epidemiology

disease occurence

part of the cl

1 from the

characteristics

graphically as P,E

patients whose

characteristics are

For the purposes
practice will be co
procedures through
this will not be e
existing techniq

which science and rationalism can be integrated into the clinical management of individual patients."

Formal Aggregation of Experience: the Clinical and Epidemiological Research Paradigms

The clinical and epidemiological research paradigms address the problem of determining and improving medical care through the formal aggregation of experience.* Although these paradigms are generally considered sub-fields of medicine and public health, their concepts and techniques can, in fact, be applied to any clinical problem solving content area. Both paradigms have evolved in the context of a clinician collecting personal and environmental characteristics (P,E) of a case, making a diagnosis (D), selecting and administering a treatment (T), and observing outcomes (O).

Epidemiology is defined by Roberts (1977, p. 16) as "the study of disease occurence in human populations." Its primary focus is the first part of the clinical problem solving task - determining the diagnosis (D) from the personal characteristics (P) and environmental characteristics (E) of the case. This focus can be represented graphically as P,E -> D. It requires the examination of patients and non patients whose various patterns of personal and environmental characteristics are compared across diagnostic groups.

----

* For the purposes of the research reported here, only accepted clinical practice will be considered. Development of new diagnostic and remedial procedures through model building, simulation, laboratory experiments, etc., will not be examined. The concern is, rather, with the optimal use of existing techniques.

By comparis

ferences perta

individual pat

1980, p. 180), ...

are of successfu

to focus the s

analyzing outcome

diagnosis [?]. Th

requires the ex

patients with va

and treatment cha

The basic

paradigms is tha

would be suppo

formal aggregat

is aggregated

characteristics

occuring group

the comparison

clinical and

i.e. empiric

It can be

representativ

The preci

of the data und

characteristics

By comparison, clinical investigation "is primarily concerned with inferences pertaining to health which can be drawn from the examination of individual patients" (Roberts, 1977, p.21). According to Friedman (1980, p.160), "The ultimate goal of clinical studies is to learn how to cure or successfully treat disease." Indeed, clinical research has as its focus the second part of the clinical problem solving task - maximizing outcome (O) by selecting a treatment (T) for a given diagnosis (D). This focus can be represented graphically as D -> T -> O. It requires the examination and comparison across outcome groups of patients with various patterns of personal, environmental, diagnostic, and treatment characteristics.

The basic tenet of the clinical and epidemiological research paradigms is that, to as great an extent as possible, clinical decisions should be supported by scientific, objective evidence based on the formal aggregation of clinical experience (Feinstein, 1977). Experience is aggregated in terms of counts of clinical problem solving characteristics and functions of these counts across various naturally occuring groups of cases. The careful recording of characteristics and the comparison of characteristics across groups is the key to both clinical and epidemiological research. This recording and comparing is simple, empirical, and does not rely on faulty and biased human memory. It can be aggregated across clinicians and cases to avoid unrepresentative or narrow samples of cases.

The precise coding of experience depends upon the characteristics of the data under investigation (Friedman, 1980, Chapter 2). Qualitative characteristics such as particular diagnostic categories tend to be compared in ter

(?). Some comm

number of cases

number of cases

in a group deve

cases in the grou

might be the

divided by the to

incidence rate

cancer patients

people in the sa

cases of cancer

individual c

to compare the

are the basis

could substanti

In addi

rates across g

significance

comparison is

attributable

different gro

particular c

recognize that

established: a

cases per

compared in terms of proportions and functions of proportions (Ipsen, 1978). Some commonly used proportions include prevalence rates (the number of cases in a group with a characteristic divided by the total number of cases in the group) and incidence rates (the number of cases in a group developing a characteristic divided by the total number of cases in the group per unit of time). An example of prevalence rates might be the number of people who have cancer in a particular area divided by the total number of people in the same area. An example of an incidence rate might be the number of people who were newly diagnosed as cancer patients in a particular area divided by the total number of people in the same area divided by the amount of time over which the new cases of cancer were discovered. These rates alone are useful to individual clinicians because they provide empirical data against which to compare their subjective impressions. Since subjective impressions are the basis of many decisions, confirming or denying their validity could substantially change an individual's decision making practice.

In addition to the incidence and prevalence rates, comparisons of rates across groups may provide useful information on the relative significance of alternate decisions (Friedman, 1980, chapter 2). This comparison is done through differences and ratios of proportions. Attributable risk is the difference of two incidence rates for two different groups; it describes the differing probabilities of a particular characteristic depending upon group membership. For example, assume that two different incidence rates for cancer had been established: a rate of 5.4 cases per 10,000 women per year and a rate of 7.8 cases per 10,000 men per year. Clearly men are at a greater risk.

The attributable

women would

year. Relative ris

ratio of two such

Quantitative

tend to be co

distributions. The

standard deviatio

single distribut

pressions. Rela

distributions acro

Of course,

same design must e

p.18) desc

a temporal seque

and undergoes a

tial state, and

state either al

by the clinica

preparation" is

tial state and

nature, the exp

The attributable risk of belonging to the group "men" compared to "women" would then be 7.8-5.4 = 2.4 additional cases per 10,000 per year. Relative risk is another means of comparing rates that is the ratio of two such incidence rates.

Quantitative characteristics, such as a particular test result, tend to be compared in terms of the characteristics of their distributions. These characteristics include the usual measures of mean, standard deviation, range, median, etc. Again, the characteristics of a single distribution are useful as a check against subjective impressions. Relative significance can be determined by comparison of distributions across groups.

Of course, coding data is not enough. As for any research model, some design must exist to guide the collection of experience. Feinstein (1977, p.18) describes the basic structure of all biomedical research as "a temporal sequence in which a preparation is exposed to a maneuver, and undergoes a response. The preparation is described according to its initial state, and the response is determined by noting the subsequent state either alone or in comparison to the initial state." Of course, for the clinical and epidemiological research paradigms, the "preparation" is a person. This person is diseased or healthy in the initial state and the subsequent state. The maneuver may be controlled by nature, the experimenter or the subject.

Research D

method by aut

field method

retrospective ve

analysis (retros

30)* Unfortuna

describe clinica

authors (Alderso

Hausen, 1978; J

Perhaps the eas

the light of the

→ D → T → D.

Epidemiolog

decision makin

can be descri

characteristics

same instant

a cross-secti

case-control s

environment

forward to a

prospective

longitudina

----

* Because c
retrospectiv
epidemiologi
these terms
because the
the terms ar

<u>Research Designs</u>. The most common dimensions of research design mentioned by authorities in the field are setting (laboratory versus field), method (observational versus experimental), data collection (retrospective versus cross-sectional versus prospective), and data analysis (retrospective versus prospective) (Fox, Hall, & Elveback, 1970)* Unfortunately, there appears to be as many different ways to describe clinical and epidemiological research designs as there are authors (Alderson, 1976; Friedman, 1980; Holland, 1970; Holland & Kahausen, 1978; Lilienfeld & Lilienfeld, 1980; Mausner & Bahn, 1974). Perhaps the easiest way to examine the design parameters is to do so in the light of the clinical decision making model described earlier: P,E -> D -> T -> O.

Epidemiological research centers around the first part of the decision making process, P,E -> D. Immediately three forms of research can be described (Lilienfeld & Lilienfeld, 1980). First, the characteristics of the case and its environment and the diagnoses at the same instant in time may all be determined. This type of study is called a cross-sectional study (or, sometimes, a prevalence study or a case-control study). Second, the characteristics of the case and its environment may be determined at a particular time and then followed forward to a diagnosis at some future time. This type of study is called prospective (or, sometimes, a cohort study or an incidence study, or a longtitudinal study, or a forward-looking study or a followup study!).

----

*. Because of the multiple use of the term prospective and retrospective, much confusion has occurred - even in the clinical and epidemiological research fields. Some authors suggest the discarding of these terms and the invention of others (Feinstein, 1977). However, because the terms do describe a useful temporal distinction and because the terms are still in common use, they will continue to be used here.

Third, the diagn

then followed

environmental cha

or, sometimes,

the confusion her

difference betwe

of personal and e

Clinical res

the second part o

diagnostic and

outcomes or afte

treatment and

outcome is viewe

The other

epidemiological

of the model,

always concerne

goal of the r

of differences

characteristics

results in an o

in an experime

to the origina

experimentally.

Third, the diagnoses of the cases may be made at a particular time and then followed backwards through existing records to previous case and environmental characteristics. These studies are called retrospective (or, sometimes, a case-control study or a case-history study). For all the confusion here, we see just one dimension, namely the temporal difference between the determination of diagnosis and the determination of personal and environmental characteristics.

Clinical research has an analogous design parameter centered around the second part of the decision making process, D -> T -> O. The diagnostic and treatment characteristics could be measured before outcomes or after outcomes (Feinstein, 1977, chapter 4). Note that treatment and outcome cannot be measured simultaneously because the outcome is viewed as a result of treatment.

The other dimension that is used to describe both clinical and epidemiological research is apparent from a closer look at the two parts of the model, P,E -> D and D -> T -> O. In either case, the research is always concerned with <u>particular</u> characteristics of the situation. The goal of the research is to relate differences in these characteristics to differences in other data. The original differences in characteristics can be defined in one of two ways: by nature, which results in an observational study, and by the researcher, which results in an experimental study. The second design dimension, then, deals with how the original differences in grouping are defined, naturally or experimentally.

Unfortunate

orthogonal, i.e.

majority of cl

in the following

1. epidemiologic

a. cross-sec

environme

same po

b. retrospec

study a

collected

environm

c. prospect

study a

characte

up until

d. prospect

randomly

characte

future p

2. clinical stu

a. retrospec

study are

Unfortunately, the two dimensions just described are not orthogonal, i.e., not all combinations are possible. Actually, the majority of clinical and epidemiological research studies can be placed in the following categories:

1. epidemiological studies

   a. cross-sectional collection of observational data - personal, environmental, and diagnostic characteristics that exist at the same point in time for the selected cases are determined

   b. retrospective collection of observational data - the cases under study are grouped on the basis of diagnosis and then data are collected from existing records on the preceding personal and environmental characteristics of interest

   c. prospective collection of observational data - the cases under study are grouped on the basis of personal and environmental characteristics of significance and then the cases are followed up until some time in the future when diagnoses are determined

   d. prospective collection of experimental data - the cases are randomly assigned to some level of personal or environmental characteristic of importance and then followed up until some future point when the diagnoses are completed

2. clinical studies

   a. retrospective collection of observational data - the cases under study are grouped on the basis of outcomes and then data are collecte

treatmen

interest

b. prospect

study are

forward

c. prospect

on rele

of treat

the outc

3. combination

these issues

Although t

ical and epi

been added in

process of phys

processes of tr

parameter

approach ignore

traditional app

w in 1853

Lilienfeld

founding tr

collected from existing records on the preceding diagnostic, treatment, personal and environmental characteristics of interest

b. prospective collection of observational data - the cases under study are grouped by naturally occurring treatments and followed forward until some future time when outcomes are determined

c. prospective collection of experimental data - the cases matched on relevant characteristics are randomly assigned to some level of treatment and then followed-up until some future point when the outcomes can be determined

3. combination studies can be run that simultaneously address many of these issues

Although the two design dimensions just described account for most clinical and epidemiological research studies, another dimension has been added in recent years. With the expanded emphasis on the thinking process of physicians, the effect of the analyzed data on the thinking processes of the practicing physician is a third design parameter. This third parameter leads to two different approaches to research. The first approach ignores the thinking process of the clinician. It is the traditional approach represented by research from Graunt in 1662 and Snow in 1853 to Collen in 1973 and the Center for Disease Control in 1977 (Lilienfeld & Lilienfeld, 1977; Collen, 1973). It consists of (1) counting the occurrences of personal and environmental characteristics,

groups of ca

3 interpreting

comparison study

study assumes th

making based on

(9).

In recent y

with how the

introduction of

represented by

(counting th

characteristics

cases, (2) comp

(4) feeding th

clinicians, (5)

(6) publishing t

These two

concerned with

differences in

two types of

process of the

in the second

fed back to the

characteristics, diagnoses, treatments, and outcomes across different groups of cases, (2) comparing these counts across groups, (3) interpreting the results, and (4) publishing the results of the comparison study in appropriate professional journals. This type of study assumes that conscientious practicioners will adapt their decision making based on the relative merit of decision alternatives (Roberts, 1978).

In recent years the second approach that has arisen is concerned with how the thinking processes of the clinician are affected by the introduction of the analyzed data. This second type of study is represented by research like Wilson et al (1975). It consists of (1) counting the occurrences of personal and environmental characteristics, diagnoses, treatments and outcomes across groups of cases, (2) comparing these counts, (3) interpreting the comparisons, (4) feeding the interpreted results back to individual practicing clinicians, (5) observing the change in clinician decision making, and (6) publishing the results.

These two approaches are clearly very similar. Both are primarily concerned with how differences in case and treatment are related to differences in diagnosis and outcome. The major difference between the two types of studies is the added emphasis on the decision making process of the physician as a systematically observed dependent variable in the second application. The aggregated integrated empirical data are fed back to the practicing clinician and changes in his/her decision aking are obser

The distin

should be clarif

Epidemiolog

amines the

characteristics a

of studies are c

First, let us tu

CROSS-SECT

In 196 ,

epidemiological

cigarette smoki

p.93).

interview and p

summarized in

Reviews of

that the majo

remains true

identified am

those

making are observed and included as part of the research.*

The distinction between clinical and epidemiological research should be clarified with the following examples.

Epidemiological Research Examples. Epidemiological research examines the relationships among personal and environmental characteristics and diagnostic categorizations (P,E -> D). Three types of studies are common: cross-sectional, retrospective, and prospective. First, let us turn to the simplest study, the cross-sectional research.

CROSS-SECTIONAL EPIDEMIOLOGICAL RESEARCH EXAMPLE

In 1961, Ferris and Anderson carried out a cross-sectional epidemiological study of chronic respiratory disease in relation to cigarette smoking and air pollution in Berlin, New Hampshire (Friedman, 1980, p.93). Data was collected from about 1200 subjects through interview and physical exam. Some of the results of the study are summarized in Table 4.

----

* Reviews of recent epidemiological and clinical research journals show that the majority of studies are of the former variety. This probably remains true because it is assumed that if a reasonable decision can be identified among a set of alternatives, the conscientious clinician will choose it.

As

Smoking Characte

Never smoked

Ex-smokers

1-10 cigarettes

11-20 cigarettes

21-30 cigarettes

31-40 cigarettes

41 or more cigare

Although the data

not followed fo

followed backward

association bet

that this study

of a personal or

categorization

because the sm

cases at the sam

RETROSPECT

After the

common design

was conducted f

myocardial inf

enfeld, 193

itted from

ilar group

data are given

Table 4
Association of Smoking and Chronic Bronchitis

Smoking Characteristic	Prevalence of Chronic Bronchitis
never smoked	15.0%
ex-smokers	18.9%
1-10 cigarettes/day	29.8%
11-20 cigarettes/day	34.2%
21-30 cigarettes/day	42.3%
31-40 cigarettes/day	61.1%
41 or more cigarettes/day	75.3%

Although the data are not related by time, i.e., particular smokers were not followed forwards to bronchitis and bronchitis sufferers were not followed backwards to smoking habits, the data seem to indicate a strong association between smoking and incidence of chronic bronchitis. Note that this study was epidemiological because it examined the relationship of a personal or environmental characteristic (smoking) and a diagnostic categorization (chronic bronchitis). The study was cross-sectional because the smoking and bronchitis data were all characteristic of the cases at the same instant in time.

RETROSPECTIVE EPIDEMIOLOGICAL RESEARCH EXAMPLE

After the cross-sectional form of epidemiologic study, the next common design is retrospective. A retrospective, epidemiological study was conducted from 1968 - 1972 concerning the relationship between myocardial infarction and use of oral contraceptives (Lilienfeld & Lilienfeld, 1980). Data on the previous use of oral contraceptives were collected from a group of 58 women treated for myocardial infarction and a similar group of 166 women without such a diagnosis. The resulting data are given in Table 5.

Association of

Group of women w
of myocardial in

Group of women w
diagnosis of myo
infarction

Clearly, the pe
the group wi
myocardial infa
because it exa
characteristic
infarction).
preexisting di
backwards in t

PROSPECT

The las
the prospecti
designed to
antibodies in
Lilienfeld
of EBV antibo
Kashman. The
etching for
resulting dat

Table 5
Association of Myocardial Infarction and use of Oral Contraceptives

	Used oral contraceptives	Never used oral contraceptives
Group of women with diagnosis of myocardial infarction	40% (23 of 58)	60% (35 of 58)
Group of women with no diagnosis of myocardial infarction	20% (34 of 166)	80% (132 of 166)

Clearly, the percentage of oral contraceptive users was much higher for the group with myocardial infarction (40%) than for the group without myocardial infarction (20%). Note that this study was epidemiological because it examined the relationship between a personal or environmental characteristic (use of contraceptives) and a diagnosis (myocardial infarction). The study was retrospective because the women were grouped on existing diagnoses of myocardial infarction or not and followed backwards in time to a previous use of oral contraceptives.

PROSPECTIVE EPIDEMIOLOGICAL RESEARCH EXAMPLE

The last common form of epidemiological research is prospective. One prospective epidemiological study conducted at Yale University was designed to examine the relationship between Epstein-Barr Virus (EBV) antibodies in the blood and the occurrence of infectious mononucleosis (Lilienfeld & Lilienfeld, 1980, p.236). Data on the presence or absence of EBV antibodies in the blood were collected on a group of 362 incoming freshman. They were then followed for four years by researchers who were watching for subsequent development of infectious mononucleosis. The resulting data are given in Table 6.

Assoc

EB virus presen

EB virus absen

Clearly, the a

susceptibility,

seroepidemiolog

personal or e

antibodies in t

mononucleosis,

grouped on the

antibodies in

mononucleosis (

Clinical

relationship

studies are us

capable of re

RETROSPEC

A retros

mortality rate

hypertension

Table 6
Association of EBV Virus and Incidence of Mononucleosis

	Developed mononucleosis within 4 years	Did not develop mononucleosis within 4 years
EBV Virus present	0% (0 of 94)	100% (94 of 94)
EBV Virus absent	15% (40 of 268)	85% (228 of 268)

Clearly, the absence of EBV antibodies was associated with susceptibility, the presence of EBV with immunity. Note that this study was epidemiological because it examined the relationship between a personal or environmental characteristic (presence or not of the EBV antibodies in the blood) and a diagnostic categorization (infectious mononucleosis). The study was prospective because the freshmen were grouped on the personal characteristic of the presence or absence of EBV antibodies in the blood and followed forward to the future diagnosis of mononucleosis or not mononucleosis.

<u>Clinical</u> <u>Research</u> <u>Examples</u>. Clinical research examines the relationship between treatment and outcome (T -> O). Two types of studies are used: the retrospective and prospective. Consider first an example of retrospective clinical research.

RETROSPECTIVE CLINICAL RESEARCH EXAMPLE

A retrospective clinical study comes from work by Childs comparing mortality rates of two surgical procedures for alleviating portal hypertension (Burdette & Gehan, 1970, p.14). Two types of portacaval shunts were posi

data collected

statistical

Asso

ted

ved

mortality

however, on fur

surgical proce

was functioning

procedures for

Association

mortality with

mortality with
damage

The end-to-si

liver damage

that this st

between the

shunts were possible, one being side-to-side, the other end-to-side. The data collected relating mortality and treatment indicated that there was no statistical difference in outcomes between the two procedures (Table 7).

Table 7
Association of Mortality and Surgical Procedure

	End-to-side shunt	Side-to-side shunt
Died	22	7
Lived	107	30
Mortality	17%	19%

However, on further analysis, the data seemed to indicate that the surgical procedure of choice depended upon how well the patient's liver was functioning. Table 8 describes the relative mortality of the two procedures for differing levels of patient liver functioning.

Table 8
Association of Mortality and Surgical Procedure Stratified by Liver Function

	End-to-side shunt	Side-to-side shunt
Mortality with minimal liver damage	4%	9%
Mortality with advanced liver damage	53%	36%

The end-to-side procedure was preferred when there was only minimal liver damage; the side-to-side procedure was preferred otherwise. Note that this study was clinical because it examined the relationship between treatment (two methods of surgery for portal hypertension) and outcome (morta

subjects with

backwards to th

surgical shunt.

PROSPECT VA

The secon

Prospective c

relationship be

patient morta

pressure betwee

and a control

data collected

Associat

Treated with o

Treated with p

The relative

of 1.06 to 1.

was 55.7% m

was 5.5 more

critical beca

(placebo) an

the patients

during therapy

outcome (mortality). The study was retrospective because it grouped the subjects with known outcomes of life or death and followed them backwards to the treatment they received of end-to-side or side-to-side surgical shunt.

PROSPECTIVE CLINICAL RESEARCH EXAMPLE

The second common form of clinical research is prospective. Prospective clinical research is demonstrated by a study of the relationship between different treatments for moderate hypertension and patient mortality (Roberts, 1977). Veterans with diastolic blood pressure between 90 and 114 were randomly allocated to a treatment group and a control group. They were followed for approximately 3.3 years. The data collected at that point are presented in Table 9.

Table 9
Association of Drug Treatment and Mortality for Hypertension

	Dead within 3.3 years	Alive after 3.3 years
Treated with drugs	4.3% (8 of 186)	95.7% (178 of 186)
Treated with placebo	9.8% (19 of 194)	90.2% (175 of 194)

The relative risk of no treatment to treatment was just 95.7% to 90.2% or 1.06 to 1. That does not seem large. However, the attributable risk was 95.7% minus 90.2% or 5.5%. For every 100 patients, the treatment kept 5.5 more of them alive after 3.3 years. Note that this study was clinical because it examined the relationship between a treatment (drug or placebo) and outcome (mortality). The study was prospective because the patients were grouped based on treatment for moderate hypertension of drug therapy or placebo and then followed forward through time to the outcome result

These exa

epidemiological

outcome; epide

characteristics

between the re

looks backwards

factors upon wh

One last

research parad

patho-physiolog

various body p

to be made. Alt

environmental

and outcomes (O

the area. It

likely or impor

model of pro

relevant data

lean body is o

The model o

today, the tra

practicing physi

the most part

teaching model

outcome resulting from treatment, death or life.

These examples typify the differences among clinical and epidemiological studies. Clinical studies focus on treatment and outcome; epidemiological studies on personal and environmental characteristics and diagnoses. They further demonstrate the difference between the retrospective and prospective approach. The retrospective looks backwards from the results of clinical decisions toward the factors upon which they were based; the prospective looks forward.

One last characteristic of the clinical and epidemiological research paradigms deserves comment. Both areas rely heavily on the patho-physiological model of the human body which specifies how the various body parts interact and defines the most important observations to be made. Although the research paradigms focus on the personal and environmental characteristics (P,E), the diagnosis (D), treatment (T), and outcomes (O), an unlimited amount of information is available on any one area. It is the medical model of the body that specifies what are likely or important effectors of diagnosis and outcome. If not for this "model of process," the research paradigms might collect totally irrelevant data. The adequacy of the model in representing the true human body is of prime importance.

The model of the human body is also important for another reason, namely, the transmission of any knowledge gained from research to other practicing physicians. Fortunately for medicine, most physicians agree with most parts of one model. Therefore, research that is based on the prevailing model of the human body can have great impact since the general framework

framework, integ

imagine how much

formally aggregat

deal of the data

When viewe

perspective, it

difficult. The fe

patterns in the

encounters. In

the particular

settings the a

diagnosis, the

performance dat

capacity of the

extremely power

experience are a

Regardless

research paradi

methodological

medicine)... ca

patients" (Robe

general framework for the information is already set. Without such a framework, integration of research would be much more problematic. Imagine how much more difficult it would be to convince physicians that formally aggregated experience should alter clinical practice if a great deal of the data they believed relevant was not collected.

When viewed from the clinical and epidemiological research perspective, it is easy to see why informal learning from experience is difficult. The learning that occurs is essentially the identification of patterns in the data collected over large numbers of clinical encounters. In medicine, the amount of data that could be collected in one particular clinical problem solving setting is large; across settings the amount of data is almost limitless. Even in reading diagnosis, the number of environmental, demographic and reading performance data is quite large. In either case, it is well beyond the capacity of the human mind's processing capacity unless the effect is extremely powerful and obvious (indeed, informal means of learning from experience are adequate for effects that are very large).

Regardless of the research design, the clinical and epidemiological research paradigms provide as noted previously "possibly the only methodological framework within which ...(science and rationalism in medicine)... can be applied to the clinical management of individual patients" (Roberts, 1978, p.174).

The Clinic

Care. The clinic

methods for tea

response to the

and integrating

improving clinica

The success

improving clinica

medicine in bo

research. This s

untreated prob

These successes

Historical

efficacy of sm

demonstrated

treatment. Budd

in the spread

of the typical a

paradigm – coun

reporting.

The research

...enfeld and

1910, as well

over 100 years a

to very diffe

The Clinical and Epidemiological Paradigms and Improved Clinical Care. The clinical and epidemiological research paradigms are formal methods for learning from experience. They are the medical field's response to the problematic nature of informal methods for accumulating and integrating experience. They have proven to be quite effective in improving clinical problem solving in medicine.

The success of formally aggregated and integrated empirical data in improving clinical care has been demonstrated often in the field of medicine in both the research areas of epidemiology and clinical research. This success has been measured against the natural history of untreated problems and the previously defined treatments of choice. These successes are documented in the following research reports.

Historical studies abound. In 1760, Bernoulli supported the efficacy of smallpox vaccinations; in 1835, Pierre-Charles Louis demonstrated the inadvisability of blood-letting as a means of treatment. Budd conducted a classic study implicating contaminated water in the spread of typhoid fever in 1873. All of this research was based on the typical application of the clinical and epidemiological research paradigm - counting, comparing across groups, interpreting results, and reporting.

The research paradigm continues to be as useful today. According to Lillienfeld and Lillienfeld (1977), "the basic reasoning processes and logic, as well as many of the methods used today, were developed well over 100 years ago. Today's epidemiological problems and situations are not very different from those encountered in the last century prior to the Bacteriolog

used the same m

sma pox vacc n

Another

epidemiological

Care Program (C

since 1964 an

detection of me

have had multi p

many ways. For

the cooccurr

characteristics

diagnosis and

Friedman, Colle

generally bee

to the patient.

Turning

experience is a

Since 1885 t

their persona

characteristic

not to many goo

to provide a

supposed effect

the Bacteriological Era." Indeed, in 1977 the Center for Disease Control used the same methods to justify measles innoculations as was used with smallpox vaccinations 125 years earlier.

Another recent research series based on the clinical and epidemiological paradigm is centered in the Kaiser - Permanente Medical Care Program (Collen, (ed.), 1978). This program has been in operation since 1964 and provides a Multiphasic Health Testing Service for early detection of medical problems. An average of over 50,000 people a year have had multiphasic checkups. This extensive data base has been used in many ways. For example, it has provided base line data for estimating the co-occurrences of many patient, diagnostic, and outcome characteristics. This base line data has been used in computer aided diagnosis and in the identification of high risk patients (Siegelaub, Friedman, Collen, & Kodlin, 1978). The results of using the system have generally been a great reduction in cost with no reduction in benefits to the patient.

Turning to the clinical area, another large data base of clinical experience is available at the Mayo Clinic (Kurland & Molgaard, 1981). Since 1885 the Mayo Clinic has kept careful records of patients and their personal, environmental, diagnostic, treatment, and outcome characteristics. This formally aggregated empirical data base has been put to many good uses throughout its history. Most recently it was used to provide a base rate of the Guillain-Barre syndrome so that the supposed effects of the swine flu vaccinations could be determined.

Clinical rese

studies performed

(?), has been co

diagnosis of many d

been aggregated acr

decision support

computer aided diag

accuracy of 91.(?)

and computer-aided

of associations

Leaper, Horrocks,

decisions such a

abdominal pain, c

to the most se

only did the erro

quality of dec

distinctly impro

the decision e

assistance were

proportion of

Unfortunately,

declined back

Staniland, & Mc

of empirical

decision alter

without the fo

Clinical research has been quite successful elsewhere. One set of studies performed at the University of Leeds, Leeds, England (deDombal, 1974), has been concerned with the use of empirical data in the diagnosis of many problems in internal medicine. Clinical experience has been aggregated across cases and made available through a computerized decision support system. In certain areas, like acute abdominal pain, computer aided diagnoses based on empirical data has achieved diagnostic accuracy of 91.1%. This is in comparison to unaided clinicians (79.7%) and computer-aided diagnosis based on clinicians' subjective estimates of associations of patient characteristics and diagnoses (82.2%) (Leaper, Horrocks, Staniland, & DeDombal, 1972). For some specific decisions such as discriminating between appendicitis and non-specific abdominal pain, computer-aided diagnosis reduced decision error from 13% by the most senior unaided clinician to 2% with computer support. Not only did the error rate decline when computer support was used, but the quality of decision making by the clinicians using the system was distinctly improved. During the time in which they used the computer, the decision errors in the clinicians' diagnoses without computer assistance were lessened by 3-4% across all patients. In addition, the proportion of appendices that perforated was reduced from 36% to 4%. Unfortunately, when the computer aid was removed physician performance declined back to the original levels (DeDombal, Leaper, Horrocks, Staniland, & McCann, 1974). Clearly, the computer support system based on empirical relationships in the data was able to identify reasonable decision alternatives more effectively than the human clinician could without the formally aggregated data.

These stu

research has con

care. In fact,

of preventive me

improved surgic

research that an

20th century (Co

Based on th

empirical data

epidemiological

the effectivene

mode of clinica

evolving field to

applied, the fie

Diagnosis, Treat

The Nature

the treatment of

the discussion

medical model (w

directly to re

four most frequ

medicine in tha

healthy reading.

These studies demonstrate that clinical and epidemiological research has contributed significantly to the improvement of medical care. In fact, for the world population as a whole, it is the knowledge of preventive medicine gained from epidemiology and the knowledge of improved surgical and pharmacuetical techniques gained from clinical research that are the main sources of increased life expectancy in the 20th century (Coleman, 1977).

Based on these data, it appears that the formal aggregation of empirical data through the application of the clinical and epidemiological research paradigms can be used to determine and improve the effectiveness of clinical problem solving according to a particular model of clinical reasoning. Now let us review the clinical problem solving field to which the clinical and epidemiological paradigm will be applied, the field of reading diagnosis.

Diagnosis, Treatment, and Outcomes in Reading

The Nature of Diagnosis and Treatment in Reading. If the diagnosis and treatment of reading problems were totally clinical in nature, then the discussion here could be quite brief. Most of the concepts from the medical model (which is itself primarily clinical in nature) would apply directly to reading. Unfortunately, reading diagnosis and treatment occur most frequently in group settings. In addition, reading is unlike medicine in that there is no naturally occurring condition called "healthy reading." A problem reader is generally not someone who has gone from *rea t

whose reading a

these character

But before look

setting for the

The clini

chapter detaile

case. During

environmental c

diagnosis, sug

outcomes are ob

in order to imp

The Curre

reading, the pe

either formali

reading invento

are primarily

tests for the g

measures.

Once a dat

work, the cli

making a diagno

diagnostic cate

gone from "healthy" adult reading to "poor" reading, but, rather, is one whose reading ability has not progressed as fast as "expected." Both of these characteristics complicate the reading situation significantly. But before looking at the complexity, let us first consider an idealized setting for the clinical diagnosis and treatment of reading problems.

The clinical problem solving model described earlier in this chapter detailed an interation between a single clinician and a single case. During this interaction the clinician collects personal and environmental characteristics and performances in order to make a diagnosis, suggest and carry out treatment, and observe outcomes. As outcomes are observed, it is possible to recycle through these decisions in order to improve the diagnosis, treatment, and outcome.

The Current State of Diagnosis and Treatment in Reading. In reading, the personal and environmental characteristics are collected either formally with standardized tests or informally with an informal reading inventory based on school texts (Spache, 1976). The tradeoffs are primarily the reported precision and reliability of standardized tests for the greater validity of the tasks one can setup in informal measures.

Once a data base of case information is available from which to work, the clinician can formalize his/her diagnostic impressions by writing a diagnosis (Ekwall, 1976; note that the question of when diagnostic categorizations are made is ignored here). Ideally, as described earl

diagnosis should

possible.

For the d

decisions appe

observational s

reading clinic

diagnoses. Eve

child twice

Kinsonnaler e

80; VanRoek

samples of e

specialists, c

from the revie

earlier in

defining the

from experien

These re

fact, readin

identical to

occasionally

problem solv

inappropriate

the diagnost

diagnosis l

clinician in

described earlier, the purpose of diagnosis is to determine cause. Diagnosis should be reliable so that valid aggregation of experience is possible.

For the diagnosis of reading problems, reliability of diagnostic decisions appears to be quite low. The results from a series of seven observational studies run over five years indicate that different reading clinicians diagnosing the same child arrive at totally unrelated diagnoses. Even more disturbing, the same clinician diagnosing the same child twice gives very different diagnoses on the two occasions (Vinsonhaler et al, 1982; Gil, 1979; Hoffmeyer, 1980; Stratoudakis, 1980; VanRoekel, 1982; Weinshank, 1980). These results held across many samples of educators who dealt with problem readers (reading specialists, classroom teachers, and learning disabilities specialists). From the review of the clinical and epidemiological paradigms given earlier in this chapter, it becomes clear that such unreliability in defining the state and cause of reading abilities makes valid learning from experience very difficult.

These results on reliability are distressing considering that, in fact, reading specialists are asked to perform tasks essentially identical to those posed in the studies above. The one argument that is occasionally put forth to belie these data is that the medical clinical problem solving model should not be applied to reading; it is inappropriate there. In a typical medical situation, a physician makes the diagnostic and treatment decisions at one point in time. The diagnosis leads directly and differentially to treatment. In reading, a clinician interacts with the case over an extended period of time, possibly making a

Even what diagno

has stated that:

reading in the la

of integration be

seems that thes

between whom th

incoordination

phenomenon, as w

show that the

the unreliabilit

be of concern

opinions of v

...nsonnaier an

indicates that

is that remed

central to

demonstrated by

Once the

administered.

of their variat

they vary incl

the differen

student behav

orientation, an

of all these

possibly making a traditional "medicine-like" diagnosis unnecessary. Even what diagnoses there are do not lead anywhere. Spache (1976, p. 436) stated that: "Despite refinements in diagnosis and remediation in reading in the last twenty or so years, there is still a widespread lack of integration between these two processes .... In many instances, it seems that these two processes are carried on by different persons between whom there is a distinct lack of communication .... the incoordination between diagnosis and remediation is a relatively common phenomenon, as we shall attempt to show." These differences are meant to show that the diagnostic perspective is inappropriate for reading and the unreliability of clinicians as they make such decisions should not be of concern. However, this argument is in direct opposition to the opinions of virtually all the experts in the field of reading. Vinsonhaler and colleagues (1982) cite study after study that clearly indicates that the cardinal principle of reading literature and practice is that remediation must be based on diagnosis. Reliable diagnosis is central to reading instruction. Therefore, the unreliability demonstrated by the practitioners in reading is a major concern.

Once the diagnosis is complete, treatments are chosen and administered. The treatments for reading problems are almost limitless in their variation (Spache, 1976). A few of the dimensions along which they vary include (1) the assumed sequence of skills needed for reading, (2) the different forms of technology applied, (3) the methods for shaping student behavior, (4) material content, (5) individual or group orientation, and so on. Any treatment, then, combines various elements on all these dimensions and more in a way that yields the wide variety of reading treatm

The last part

determination of

treatment. Accord

Practically all

utilize gain on rea

reasonable criter

basis for the o

resolved, the pe

reflect this.*

Reading Ins

clinical problem

Duffy, 1981;

professional rea

problem solvin

occurs in a gro

student being

other students

substantially

The prima

realities of

---

*Many of the
a student mig
measurement c
al., 1979; C

of reading treatment possibilities.

The last part in the clinical problem solving process is the determination of outcome or change that occurs after the introduction of treatment. According to Spache (1976, p. 273), variation (Spache, 1976). "Practically all reading programs, methods of reading, and experiments utilize gain on reading tests as their criteria of success." These are reasonable criteria. The student performances on some tasks were the basis for the original diagnosis. As the student's problems are resolved, the performance on these same measures should change to reflect this.*

<u>Reading Instruction in the Classroom</u>. Now, into this manageable clinical problem solving environment, the realities are introduced (Duffy, 1981; Jackson, 1968). The fact of the matter is that little professional reading instruction is conducted in the idealized clinical problem solving setting. Instead, most reading diagnosis and treatment occurs in a group setting. Besides the clinician and the particular student being helped, many other students are also in the room. These other students add a degree of complexity to the situation that may substantially alter the characteristics of the instructional setting.

The primary difference between the ideal clinical setting and the realities of classroom settings is the number of interpersonal

----

* Many of the measurement tasks for reading include connected text that a student might easily remember. Therefore, alternate forms of measurement devices may often be required to determine outcome (Borg & Gall, 1979; Calfee & Drum, 1979).

interactions tha

there is a one-on

the student, th

many-on-many inte

divide up his,h

diagnosis, treatm

student-to-studen

interfere with o

Because they e

treatments that

In general, as

of a clinician's

of inter-stude

individualized

clinician's res

and a standar

management tech

a clinician c

one-on-one giv

uniformity in

and real-life

1981) indicat

commercially

feedback, "it

to be cogn

differential

traditionally-dev

interactions that can be taking place. In the ideal clinical setting, there is a one-on-one interaction. The clinician can directly respond to the student, the student to the clinician. The classroom setting has a many-on-many interaction. At the first level the clinician must somehow divide up his/her energies across many students, each requiring a diagnosis, treatment and followup. Just as important, however, are the student-to-student interactions. There inter-student relations can interfere with other clinician-student interactions or facilitate them. Because they exist, the inter-student interactions change the range of treatments that the clinician can administer to the individual students. In general, as the number of students in a class increases, the amount of a clinician's energies that must be spent on management and control of inter-student interactions increases and the amount spent on individualized treatment decreases. At some point, little of the clinician's resources are spent on any type of individualized treatment and a standardized, non-discriminatory treatment with accompanying management techniques is the rule. Clearly, the range of treatments that a clinician can apply are greatly attenuated by classroom size: one-on-one gives total individuality in treatment; one-on-many gives uniformity in treatment across differing problems; one-on-few may be a good real-life compromise. At this time, research reported by Duffy (1981) indicates that most teachers generally take students through commercially prepared materials providing only corrective and affective feedback. "It is difficult for them to remember that they are supposed to be cognitive information processors who make individualized differential instructional decisions on the basis of rationally-developed hypotheses (p. 11)."

This reality

epidemiological pa

place, the con

treatments difficu

energy, and stud

than individualize

generally cognitiv

to interfere with

varies over diffe

of treatments tha

to make substanti

If these co

group settings

receive in the

clinical and

differential tr

are extremely

research in ins

disentangling t

Factors

Interpreting t

Bloom, 1976;

Haren, D'shaw,

learning tha

This reality of reading instruction makes applying the clinical and epidemiological paradigms to reading more problematic. In the first place, the constraints of classroom activity make differential treatments difficult. The clinician has a limited amount of time, energy, and student cooperation and usually, therefore, provides less than individualized attention. In addition, since reading treatments are generally cognitive in nature, the social demands of other students tend to interfere with what treatment there is. Finally, as classroom size varies over different clinicians and different settings, the total range of treatments that are applied increases, requiring more data collection to make substantive comment.

If these complications are not enough, the added complexities of group settings make it difficult to define what treatment the students receive in the first place (Jackson, 1968; Good & Brophy, 1973). Clinical and epidemiological research require reliable diagnosis and differential treatments. But the factors that make treatments different are extremely interrelated and difficult to assess. Fortunately, research in instruction and reading have provided several models for disentangling this complexity.

Factors Affecting Reading Instruction. Several models exist for interpreting the complexities of the classroom and reading instruction (Bloom, 1976; Carroll, 1963; Gagne, 1977; Fisher, Filby, Marliave, Cahen, Dishaw, Moore, & Berliner, 1978). One model of individual learning that suggests the significant aspects of treatment and has had a great deal of

The Carroll mode

improving a part

a student to impro

for instruction a

student actually

understand the

quality of instru

these factors is

The complex

models. One by

treatment in gro

characteristics

... abilities

factors can s

about the abili

reading. But,

for interpretin

provide a fram

of actual instr

Measuring

treatment, any

measure the mos

measurement of

a great deal of influence in education is the model by Carroll (1963). The Carroll model defines five factors that affect treatment aimed at improving a particular task. These are the amount of time it would take a student to improve under ideal conditions (aptitude), the time allowed for instruction by the clinician, the part of the allowed time that the student actually uses (perseverance), the student's ability to understand the treatment tasks in which s/he is to engage, and the quality of instruction compared to some theoretical ideal. Variance in these factors is variance in treatment.

The complexities of the classroom are described by other heuristic models. One by Dunkin and Biddle (1974) details the factors that affect treatment in group settings. These factors include everything from characteristics of the clinician and the group environment to entering pupil abilities and classroom activities and materials. So many, many factors can supposedly affect the treatment that it raises questions about the ability to define differentially effective treatments in reading. But, of course, the heuristic models not only provide a means for interpreting the complexities of classroom treatment, they also provide a framework for observing and recording the significant aspects of actual instruction.

Measuring Reading Instruction. With the great complexities of treatment, any clinical research in reading must be very careful to measure the most appropriate factors in treatment. Coding forms for the measurement of group instruction abound. Simon and Boyer (1967, 1972) have developed

and dozens of

alternatives te

measure instruct

difficulty of

Others, like the

1973 emphasize

many ways, yet

to the studen

have apparently

Possibly this

significantly wi

Walberg point

the threshold c

has been inte

has been vigor

of informatio

and epidemiolo

among studer

differences in

This ch

that original

Sieman (196

when they use

together mor

Brophy & Eve

have developed a catalog of observational instruments containing dozens and dozens of different alternatives. Unfortunately, the dozens of alternatives tend to be merely variations on a theme. Some instruments measure instruction in terms of clinician leadership style, enthusiasm, difficulty of questioning, etc. (Brophy & Evertson, 1976, p.174). Others, like the Brophy & Good Dyadic Observation Scales (Good & Brophy, 1973) emphasize the quality of the teacher-student interaction in so many ways, yet fail to examine how the content of instruction is related to the student's instructional deficits. The realities of a classroom have apparently focussed most of the attention on non-content concerns. Possibly this is most appropriate because content does not vary significanly within one instructional group. However, as Peterson and Walberg point out (1979, p.183) "Research on teaching appears to be at the threshold of fundamental conceptual reorganization .... First, there has been intensified interest in student variables .... Second, there has been vigorous study of cognitive and individual differences within an information-processing framework." This is precisely what clinical and epidemiological research would suggest. The individual differences among students, i.e., differences in diagnosis, should lead to differences in treatment and outcome.

This change in emphasis is spurred on by some of the same factors that originally turned the research focus to teachers. Studies by Coleman (1966) and others suggested that teachers made no difference when they used "school" as the unit of analysis. Clearly, "school" lumps together more and less effective teachers to give uncertain results (Brophy & Evertson, 1976, p. 9). The clinical and epidemiological paradigms sugges

as the unit of

faster and slower

measures of treat

All these co

potentially diff

the concerns of

yet draw on a

instruments. Thu

problematic tha

problems with t

process. As S

observation ins

learn about c

classroom?" The

Models of

treatment in

far, namely,

activity is

decifies the

along the da

Unfortunately,

difficult.

paradigms suggest that looking at individual teacher and using "class" as the unit of analysis is also too gross a level because it combines faster and slower students to give mixed results, i.e., inadequate measures of treatment effectiveness.

All these concerns combine together to make treatment observation a potentially difficult task. Observation must attend simultaneously to the concerns of the Carroll model, Dunkin and Biddle's heuristic, and yet draw on a relatively homogeneous and narrow set of observational instruments. Thus, clinical research in reading may be much more problematic than epidemiological research. The mechanism by which these problems with treatment observation are alleviated is the model of case process. As Stallings (1977, p.19) points out, "Before choosing an observation instrument, you must answer the question: What do I want to learn about children, myself, or other adults participating in the classroom?" The model of case process provides the answer.

<u>Models of the Reading Process</u>. As the activities of diagnosis and treatment in reading have been explored, one issue has been ignored so far, namely, the model of case process. Clinical problem solving activity is centered around a model of how the case works. This model specifies the data that is significant and the expected relationships among the data. Such a model of the reading process is needed. Unfortunately, locating a model that is widely accepted could well be difficult.

Chall (198

proliferation of

best way to teach

on the answer," p

field of great

1977). Yet the con

epistemological

accepted and stab

of the human body

medical data and

dissemination of

across physicians

competing model

research and pra

compromise would

however, quoting

strong emotion

specialists, t

researchers. T

characteristic

p. 7).

The Mode

controversial

nevertheless

Chall (1967) highlighted the problems resulting from the proliferation of various reading models with her comment "What is the best way to teach a young child to read? No two people, it seems, agree on the answer." (p. 1). Apparently there are many competitors in the field of creating models of reading (Singer & Ruddell, 1976; Davis, 1971). Yet the cornerstone upon which the successes of clinical and epidemiological research in medicine are built has been that of an accepted and stable model of the human body. The anatomy and physiology of the human body is a widely accepted model that identifies significant medical data and their interactions and that serves as a basis for dissemination of information to practitioners. Without it, results across physicians would be largely noncomparable. In reading, with competing models that can be substantially different, such guidance for research and practice has not been possible. One would hope that some compromise would be possible and an all encompassing model would evolve. However, quoting Chall again, "The overriding impression was one of strong emotional involvement on the part of authors, reading specialists, teachers, administrators, and, unfortunately, even researchers. Their language (defending reading models) was often more characteristic of religion and politics than of science and learning." (p. 7).

<u>The Model of Reading and Learning to Read (MORAL)</u>. However controversial the various models of the reading process are, some model must, nevertheless, be chosen to guide decisions during clinical and

epidemiological

model is the dat

actually run, th

data collection.

validity in th

valued. The mode

Read (MORAL)

Clinical Studie

Sherman, in p

This model incl

them, and in

information-pr

The orig

as a reading c

designed to as

describes a s

be able to

personal and

that would ef

overly conce

defined wha

factors that

orientation

reading is y

very least,

epidemiological research. At an initial stage, the true key to such a model is the data that it deems significant to collect. When a study is actually run, the only point at which the model is operationalized is in data collection. Therefore, a model was desired that had high face validity in the depth and range of reading related behaviors that it valued. The model chosen here was the Model of Reading and Learning to Read (MORAL) developed by Dr. George B. Sherman and expanded by the Clinical Studies project of the Institute for Research on Teaching (Sherman, in progress; Vinsonhaler, Weinshank, Cureton, & Blatt, 1980). This model includes the definition of processes, factors that affect them, and instruments for measuring them: the requirements for a good information-processing model (Calfee, 1981).

The original model defined by Sherman was the result of many years as a reading diagnostician and college professor of reading. The model, designed to assist with the diagnosis and treatment of reading problems, describes a series of Critical Performances that a "healthy" reader must be able to demonstrate together with the concurrent cognitive skills, personal and environmental factors, learning history and learning skills that would effect the state of the Critical Performances. Rather than be overly concerned with the way learning to read took place, Sherman defined what a good reader should be able to do and then watched for factors that seemed to affect these abilities. One result of this orientation is that the face validity of the Critical Performances for reading is very high. Research that collects these data will, at the very least, be collecting much important performance data.

The Clini

research in the

techniques for

The Clinical St

processor of

conceptualized

2) processes th

and 3) produces

both. A parti

receives as inp

with his past

way. Outputs ar

memory) and s

reading task as

Six readi

the health of a

inadequate, sc

deficient. The

1. Instant wor

recognize

associate t

in the mind

2. Decoded wo

recognize a

sound symbo

The Clinical Studies Project expanded the model with relevant research in the field of reading and with descriptions of measurement techniques for many of the significant data highlighted by the model. The Clinical Studies version of the MORAL conceptualized the reader as a processor of information. Any information processor can be conceptualized as an entity that 1) receives input from the environment, 2) processes this input in conjuction with its own memory of past events and 3) produces an output that impacts its memory, the environment, or both. A particular reader attempting a particular reading task, then, receives as input the requirements of the task. This input, together with his past knowledge of reading and language, are processed in some way. Outputs are produced; some are not observable (e.g., changes in memory) and some of which are observable (e.g., performance on the reading task as measured in a particular way).

Six reading and language performances are considered critical to the health of a reader. To the degree that these performances are inadequate, some tasks in the range of reading situations will be deficient. The six performances are:

1. Instant word recognition performance: defined as the ability to recognize a certain set of words instantly. (Recognize means to associate the visual stimuli of a word with its internal prototype in the mind.)

2. Decoded word recognition performance: defined as the the ability to recognize a set of words using various association strategies (e.g., sound symbol association). This recognition is not immediate and

takes more

performance

3. Vocabulary

4. Oral Readi

correctly

5. Silent Re

answer spe

6. Listening

answer spe

The MORAL

Critical Perfo

the associated

factors that

set of learnin

the significa

various readi

and motivatio

MORAL, then,

reported here

from the subj

takes more effort on the part of the reader than does instant word performance.

3. Vocabulary performance: defined as the ability to give word meanings

4. Oral Reading Performance: defined as the ability to read text aloud correctly with appropriate phrasing, fluency and intonation.

5. Silent Reading Comprehension Performance: defined as the ability to answer specific questions on text read silently.

6. Listening Comprehension Performance: defined as the ability to answer specific questions on text read aloud by someone else.

The MORAL goes much further than just the specification of the Critical Performances. For each Critical Performance the MORAL specifies the associated effecting factors, that is, the student and environmental factors that affect the performance. For each performance, in turn, a set of learning or change processes is defined. These processes describe the significant aspects of learning history in reading that affect the various reading performances. The very pertinent concerns of attention and motivation are viewed as effectors of the Critical Performances. The MORAL, then, is the model of the reader that was used in the research reported here to define the significant data that needed to be collected from the subjects in the study.

Summary of Relat

Clinical

is widely used

available meth

efficiency of c

experience, is

tasks and the

to valid inf

clinical and e

effective al

experience apa

the compariso

non-representa

have been us

them to the f

straightforwa

some poor cli

practitioners

an excellent

epidemiologic

Summary of Related Research

Clinical problem solving is a general mode of human thinking that is widely used in many content areas. Unfortunately, the generally available method for determining and improving the effectiveness and efficiency of clinical problem solving, namely informal aggregation of experience, is inadequate. The structure of the clinical decision making tasks and the biases of the human problem solver do not lend themselves to valid informal learning from experience. The formal methods of the clinical and epidemiological research paradigms seem to be a potentially effective alternative. They direct the explicit recording of clinical experience apart from the human mind's biases and faults; they emphasize the comparison of such records across cases and clinicians to avoid non-representative or narrow experience. Where these research techniques have been used in medicine, they have proven quite effective. Adapting them to the field of reading diagnosis and treatment seems to be a straightforward generalization. This field has been shown to suffer from some poor clinical problem solving performance on the part of its practitioners. With a recent model of the reader available, it becomes an excellent area in which to examine how well the clinical and epidemiological research paradigm does generalize to other fields.

Objectives

The purpose of t
research paradig
These research
diagnosis and tr
the aggregation
have proven to b
means for learn
paradigms to re
clinical exper
in problem reac

The specif

1. to determi
research pa
general, a
diagnosis a

2. to evaluate
study in
diagnosis a
epidemiologi

CHAPTER 3

Methods and Procedures

Objectives

The purpose of this thesis was to apply the clinical and epidemiological research paradigms to the field of reading diagnosis and treatment. These research paradigms were developed in the field of medical diagnosis and treatment and have served there as effective guides for the aggregation and integration of clinical experience. These techniques have proven to be conspicuously better than the traditional informal means for learning from experience. It was hoped that generalizing these paradigms to reading might point the way toward more effective use of clinical experience in reading to improve the quality of care received by problem readers.

The specific objectives were as follows:

1. to determine how to apply the medical clinical and epidemiological research paradigms to any other clinical problem solving field in general, and to carry out such a study in the field of reading diagnosis and treatment in particular.

2. to evaluate the results of the clinical and epidemiological research study in reading with respect to further understanding of reading diagnosis and treatment and the use of the clinical and epidemiological paradigms

a. to exam

and diag

1) the

vario

2) asso

chara

3) asso

char

anal

4) prob

char

b. to ex

charact

1) the

liv

2) as

wi

3) as

wi

4) p

c

a. to examine the epidemiological data of personal, environmental, and diagnostic characteristics for

1) the nature of reading diagnosis and baseline likelihoods of various characteristics with cross-sectional data analysis;

2) associations between diagnoses and personal or environmental characteristics with retrospective data analysis;

3) associations between personal or environmental characteristics and diagnoses with prospective data analysis;

4) problems with the definition or measurement of the various characteristics or the use of the epidemiological paradigm.

b. to examine the clinical data of treatment and outcome characteristics for

1) the nature of reading treatment and outcome and baseline likelihoods of various characteristics;

2) associations between outcomes and treatment characteristics with retrospective data analysis;

3) associations between treatment characteristics and outcomes with prospective data analysis;

4) problems with the definition or measurement of the various characteristics or the use of the clinical paradigm.

The remaind

first objective

epidemiological

problem solving

epidemiological

this division

use the clinic

experience. The

required by the

used in readi

paradigm outsid

Clinical

problem solving

the decisions

the first prere

health of the

background. T

practitioners

model of case

Procedures

Clinical and e

i.e., an "an

particular cli

assist in the

The remainder of this chapter describes the achievement of the first objective, that is, the determination of how the clinical and epidemiological research paradigms are generalized to any clinical problem solving field in general and how the particular clinical and epidemiological research study in reading was performed. The purpose of this division is to provide a guide for other fields that may wish to use the clinical and epidemiological paradigms for learning from experience. The description of general procedures details the techniques required by the paradigm and their importance. The specific procedures used in reading provide a precise history of one attempt to use the paradigm outside of medicine.

Clinical and epidemiological research evolved from the clinical problem solving setting, and many of its requirements are dictated by the decisions that must be made by clinicians in such settings. One of the first prerequisites is a knowledge of what is significant in the health of the case and what is not; what is foreground and what is background. This knowledge should be as widely accepted by the practitioners and researchers as possible. It can be conceptualized as a model of case process.

1. Procedures for Defining a Model of Process

Clinical and epidemiological research require a model of case process, i.e., an "anatomy and physiology" for the field to be studied. For any particular clinical problem solving field, the following procedure will assist in the creation of such a model.

Genera

the develop

accurate and

be obtainabl

1. Define t

2. Define

function

3. For ea

characte

likeliho

4. For ea

environm

Estimate

5. For eac

define s

will alt

change f

6. For each

characte

that can

and seve

7. For an

characte

<u>General</u> <u>Procedure</u>. The purpose of this procedure is to assist with the development of an operational model of case process. The more accurate and precise the description, the better. But a useful model may be obtainable even from general information. Complete the steps below:

1. Define the type of system to be diagnosed, treated, and followed up.

2. Define the set of performances that are critical to case functioning.

3. For each Critical Performance, define any concurrent case characteristics that could affect the performance. Estimate the likelihood, direction, and severity of each effect.

4. For each Critical Performance, define any past personal or environmental characteristics that could affect performance. Estimate the likelihood, direction, and severity of each effect.

5. For each Critical Performance and concurrent case characteristic, define sets of actions that can be initiated by a clinician that will alter the performance or characteristics. Estimate the expected change for each set of actions.

6. For each action that can affect Critical Performances or case characteristics, list all case, clinician, and environmental factors that can attenuate the effect. Estimate the likelihood, direction, and severity of each effect.

7. For any of the effectors of Critical Performance or case characteristic, describe the cause - effect reasoning of the

associat

5. Define e

of syste

6. Iterate

stable.

Procedu

treatment c

MOPAL crea

Studies proj

model, as de

of careful

required co

effectors, 1

Measurem

Once the mod

a series of

the model a

possible wit

the state

performances

association, if possible.

8. Define expectations of Critical Performances for different classes of systems.

9. Iterate through the previous steps until the descriptions become stable.

Procedures in Reading. For the research in reading diagnosis and treatment described here, the Model of Reading and Learning to Read (MORAL) created by Dr. George B. Sherman and expanded by the Clinical Studies project of the Institute for Research on Teaching was used. This model, as described in the related research, was created over many years of careful clinical experience in reading. It includes all of the required components, i.e., the Critical Performances and their effectors, the improvement factors and their effectors.

II. Measurement Device Creation Procedures

Once the model of case process has been defined, the next requirement is a series of measurment devices for all of the Critical Performances of the model and as many concurrent and historical case characteristics as possible within cost constraints. These measures will be used to deduce the state of the case's Critical Performances and changes in these performances after treatment.

General

the develop

defined by

1. Estimat

for inf

2. For ea

tasks t

perform

of the

for rep

reliabi

3. Similar

define

validit

4. Within

measure

many co

5. During

reliabi

Proced

of a testing

the MCPAL,

General Procedure. The purpose of this procedure is to assist with the development of measurment devices for the Critical Performances as defined by the model of case process.

1. Estimate the maximum amount of time and other resources available for information collection for each case.

2. For each Critical Performance in the model of case process, define tasks to be presented to the case that will elicit the level of case performance. If the measurement task itself can change the behavior of the case, alternate forms of each measurement should be available for repeated case measurement. Estimate the generalized cost and the reliability and validity of each measurement procedure.

3. Similarly, for each concurrent or historical case characteristic, define measurement procedures and estimate cost, reliability and validity.

4. Within the constraints of available resources, choose a set of measurement procedures that measure all Critical Performances and as many concurrent and historical effectors as possible.

5. During actual usage of the measurement procedures, estimate actual reliability if possible.

Procedures in Reading. Many person-hours were spent in the creation of a testing battery that was true to the concepts in the MORAL. Given the MORAL, measurements were needed for each of the six Critical

Performances:

vocabulary, o

comprehension

decision was

decision all

resulted in

appeared mo

many sets of

comprehensio

could be fou

graded. Fo

graded sele

of comprehe

Performances

battery ba

this specia

stimulus m

recording t

Instant

SORT is

ten lis

level of

words ra

the SOR

the exam

in which

Performances: instant word recognition, decoded word recognition, vocabulary, oral reading, silent reading comprehension, and listening comprehension. Given the difficulty of creating reading tests, the decision was made to use existing standardized tests. While this decision allowed us to start up relatively quickly and inexpensively, it resulted in less than ideal adherence to the MORAL. The single area that appeared most problematic was the reading of connected text. Although many sets of graded selections could be found, none used consistent comprehension frames as defined by the MORAL. Although many selections could be found with consistent comprehension frames, none of these were graded. For the determination of grade level reading, two tests with graded selections were chosen. This meant, however, that different types of comprehension were not independently measured. The MORAL Critical Performances were measured, then, with a specially constructed test battery based on existing tests of reading ability. The components of this special battery were as follows: (note that a sample of the test stimulus materials is included in appendix A and a sample booklet for recording the student's performance is included as appendix B)

1. Instant Word Recognition : the Slosson Oral Reading Test (SORT). The SORT is an individual test of word recognition ability composed of ten lists of twenty words each. Each successive list is at a higher level of difficulty than the one preceding it. The difficulty of the words ranges from primer to high school. The standard procedures for the SORT give the student five seconds to recognize a word before the examiner moves the student on. The student is started at a list in which s/he can recognize all 20 words and is to continue until

none of

records of

In

were empl

planned,

reorderin

were st

performan

read onl

specialis

problems

examiner

immedia

ultimate

The mis

gradatio

Finally,

last lis

language

The

recogni

1. Decoded

SGDT was

describe

student

none of the 20 words in a list is called correctly. The examiner records only right or wrong responses.

In the test battery constructed here, non standard procedures were employed. First, since a pre-, mid-, and post-test were planned, three variations of the SORT were created by randomly reordering the words in each of the ten lists. Second, all students were started in the first list because of anticipated poor performances. Third, the students were stopped when they correctly read only five or fewer words correct in a given list. The reading specialist felt that going further could well introduce affect problems in ths student's reaction to the test battery. Fourth, the examiner recorded word calls in three categories: instant (immediately called correctly), mediated (not immediate, but ultimately called correctly), and miscalled (not called correctly). The miscalls were recorded phonetically. This allowed a finer gradation of student performance than with the normal procedures. Finally, the student was asked to define the twenty words in the last list that s/he attempted. This was an attempt to see if language experience was interfering with word recognition.

The special test battery, then, measured instant word recognition using the SORT with non-standard procedures.

2. Decoded Word Recognition : the Slosson Oral Reading Test (SORT). The SORT was also used as a measure of overall decoding ability. As described earlier, the examiner identified which words, if any, the student did not recognize instantly but ultimately did call

correcti

test for

from m sc

In a

of the

Gates-Mac

test pre

23 nons

battery,

The exa

Again, b

of the t

3. Vocabula

the Spa

in the

student

SCRT (

by the

correct

Th

comprehen

includes

test c

these s

for th

correctly. The miscalled words were phonetically transcribed on the test form. Further data on overall decoding ability was available from miscues on the oral reading tasks to be described below.

In addition to these measures of overall decoding ability, some of the specific subskills of decoding were checked with the Gates-McKillop Recognizing and Blending Common Word Parts test. This test presents three lists (23 consonant blends, 23 phonograms, and 23 nonsense words) for the student to call aloud. In this test battery, each student was asked to call all 69 words or word parts. The examiner categorized the response as correct or incorrect. Again, because of the pre-, mid-, post-test plan, three variations of the test were created by randomly reordering each list.

3. Vocabulary : selected words from the Slosson Oral Reading Test and the Spache and Durrell paragraphs. Two vocabulary measures were used in the test battery. The first measure was based on the SORT. The student was asked to give the definition of all words on the last SORT list that s/he had attempted. The words were pronounced aloud by the examiner and the student's response was categorized as correct, incorrect, or no attempt to respond.

The second measure was based on the silent reading comprehension selections and the listening comprehension selections included in the student's test battery (described below). During test construction, words were chosen by the research staff from these selections. Again, the student was asked to give a definition for these words pronounced aloud by the examiner. Again, the response w

No a

defined

parameter

comprehen

attempted

The

measures

effector

reading c

4. Oral Read

Test. Th

the para

dramatica

Reading t

and used

going fur

two minus

publisher

and D wer

the MSU re

The

but these

directly t

program ava

response was categorized as correct, incorrect, or no attempt.

No alternate forms were used here because the words to be defined for a particular student depended wholly upon other design parameters and his/her reading performance (i.e., what form of comprehension measures were used and what was the last SORT list attempted).

The logic of choosing these sets of words for vocabulary measures stems from the MORAL, which views word meanings as an effector of the major reading skills of word recognition and silent reading comprehension.

4. Oral Reading : the Gray Oral Reading Test and the Sullivan Placement Test. The Gray presents a series of paragraphs to be read aloud. As the paragraphs proceed, their difficulty level rises very dramatically. The student's task was to read the paragraphs aloud. Reading time, miscues and other oral reading errors were recorded and used to compute grade equivalents. The student was stopped from going further when his/her reading time for any paragraph exceeded two minutes. Four forms of the test were available from the publisher (A, B, C, and D). In our test batteries, versions A, B, and D were used as the alternate forms (form C was not available in the MSU reading clinic files).

The Sullivan also presented short paragraphs to be read aloud, but these were substantially shorter, simpler, and were keyed directly to placement into the reading materials of the treatment program available at the subjects' school. Only one form was availabie t

test, this

5. Silent Read

and the Spac

one for

comprehensi

in similar

through di

alternate

comprehensi

Since

of student

the pre-m

forms for

Therefore,

would be p

Howev

having enc

reviewing

was made

the Durre

treat the

design,

Durrell t

be used

available but since no measure of retention was extracted from the test, this was not viewed as a problem.

5. Silent Reading Comprehension : the Durrell Silent Reading Paragraphs and the Spache paragraphs. Two comprehension measures were required, one for silent reading comprehension and one for listening comprehension. The MORAL describes these two kinds of comprehension in similar ways. They both tap into the same semantic net, but through different means. Therefore, it was desirable to have alternate forms of the same test to measure these two different comprehensions.

Since these tests are presented as connected text, the chance of student recall of previous testing materials seemed great given the pre-mid-post test design. This would require three alternate forms for each of the listening and silent comprehension measures. Therefore, ideally, six equivalent forms of comprehension selections would be provided.

However, no silent reading test using graded paragraphs and having enough alternate forms could be located. Therefore, after reviewing several standardized tests for comprehension, the decision was made to use the Spache comprehension paragraphs (two forms) and the Durrell silent reading paragraphs (two forms). It was decided to treat these tests as if they were equivalent forms. The planned design, then, called for one form of the Spache and one form of the Durrell to be used for each test session. One of these forms would be used for silent reading comprehension and one would be used for listenin

booklet

Wh

read a

questio

researc

detail,

reading

questic

selecti

proceed

pretest

sequent

changed

change

was tur

6. Listeni

the Sp

silent

selecti

it sile

7. Attentio

measure

examiner

motivatio

listening comprehension. For specifics of the design, see the test booklet design table below.

Whatever the selections used, the task for the student was to read a paragraph silently, then to answer three or four specific questions from the examiner. The questions were created by the research staff in an attempt to measure separation of main idea and detail, inference, and sequence. The examiner recorded the student's reading time, his/her response to the main idea and inference questions, and the sequence of free recall of the contents of each selection. The students started with the second grade paragraph and proceeded upward through consecutively higher selections. During pretest, a stopping criteria was defined based on the number of sequential memories and on reading time. After pretest, this was changed to ignore reading time and use number of memories only. This change was implemented when it became clear that the time criterion was turning what was intended to be a power test into a speed test.

6. Listening Comprehension : the Durrell Silent Reading Parapraphs and the Spache paragraphs. The same procedures were used as for the silent reading comprehension except that the examiner read the selection aloud to the student rather than having the student read it silently.

7. Attention/Motivation : a most critical effecting factor. No direct measure was available. Rather, through interview questions and examiner judgement during test administration, the past and present motivation of the student was estimated.

In addition

Performances,

from initial

debriefing af

The te

student and

performance.

hour and 45

materials f

three ring

experimenter

script. At e

spelled out

were genera

This guaran

materials w

stimulus

appendices

The pa

description

tasks were

defined. Co

test batter

in versions

number of

version of

In addition to these specific measures of the MORAL Critical Performances, information on all of these areas and more were obtained from initial structured interviews with the student and from examiner debriefing after the test session.

The test was conducted using a set of stimulus materials for the student and a protocol notebook on which the examiner recorded student performance. Depending on the student, it took between one hour and one hour and 45 minutes to administer the battery. The student's stimulus materials for the measurement task were kept in plastic sheets in a three ring binder. The student protocol notebook in which the experimenter entered the student's responses was formatted like a script. At each step, the next action to be taken by the examiner was spelled out. The protocol notebooks for recording student performance were generated by computer for each student for each testing session. This guaranteed proper choice of test versions for the student. All test materials were prepared on a text processing system. A sample of the stimulus materials and test protocol notebooks are included as appendices A and B, respectively.

The particular characteristics of a test were defined by one of the descriptions in Table 10. Three different combinations of isolated word tasks were defined; eight different combinations of connected text were defined. Combining these tasks into one test gave 24 variations of the test battery. The students in the study were randomly allocated to these 24 versions of the total test at pretest time in such a way that the number of students taking any version of the total test battery or any version of an individual subtest were approximately equal.

Battery number
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524

These

with some

than introd

student,

Gray, 7) the

comprehensi

vocabularie

Table 10
Test Booklet Characteristics

Battery number	SORT version	Gates version	Gray version	Silent version	Listening version
601	1	1	1	1	3
602	2	2	2	1	4
603	3	3	3	2	3
604	1	1	1	2	4
605	2	2	2	3	1
606	3	3	3	3	2
607	1	1	1	4	1
608	2	2	2	4	2
609	3	3	3	1	3
610	1	1	1	1	4
611	2	2	2	2	3
612	3	3	3	2	4
613	1	1	1	3	1
614	2	2	2	3	2
615	3	3	3	4	1
616	1	1	1	4	2
617	2	2	2	1	3
618	3	3	3	1	4
619	1	1	1	2	3
620	2	2	2	2	4
621	3	3	3	3	1
622	1	1	1	3	2
623	2	2	2	4	1
624	3	3	3	4	2

These five parameters for test booklet generation were combined with some nonvarying tasks to yield a test battery that consisted of 1) an introduction to the student, 2) a structured interview with the student, 3) the SORT, 4) the SORT vocabulary, 5) the Gates-McKillop, 6) the Gray, 7) the Sulliven, 8) a break for the student, 9) the Silent reading comprehension, 10) the Listening comprehension, 11) the comprehension vocabularies, and 12) the examiner debriefing.

ii. Proced

Given the

significant

diagnostic

and reading

achieved thr

is a writte

Another is

low inferen

another is

from many s

decision a

General

form. If th

Critical p

factors hav

written ch

guarantee

possibiliti

basis for d

would be co

The se

decision r

the previo

III. Procedures for Reliable Diagnosis

Given the model of case process and devices for measuring its significant constructs, the next requirement is reliability in diagnostic categorization. The results of previous research in medicine and reading indicated that reliabile diagnostic categorization can be achieved through the use of external memory aids. One type of memory aid is a written form to guide a clinician through the diagnostic task. Another is a simple set of computerized decision rules that would make low inference diagnoses directly from raw student performance. Still another is a complex computer system that tries to combine information from many sources to give a finer grained diagnosis. All three forms of decision aid were investigated in this study.

General Procedure. The first decision aid mentioned was a written form. If the model definition procedures were followed, a series of Critical Performances and their concurrent and historical effecting factors have been defined. One simple decision aid would be merely a written checklist of all Critical Performances and their effectors to guarantee that the clinician at least considered all of the possibilities. Other additions might include questions concerning the basis for decisions and suggestions for remedial practice. One such form would be completed for every diagnosed case.

The second decision aid would be a set of simple computerized decision rules. Given the data from the measurement devices created in the previous procedures, these rules would give diagnostic classifi

could b

would be

be prob

Fi

guarant

content

through

could t

simulat

about s

Su

Student

conside

aid,

given

classifications. A threshold for performance on each measurment task could be established. Any case that scored better than the threshold would be diagnosed as "not problematic;" a case that scored worse would be "problematic."

Finally, a complex computer simulated clinician could be used to guarantee reliable diagnoses. All of the measurements collected in the content area would have to be related to diagnostic categorization through complex and probabilistic computer programs. These relationships could then be used to define the memory of a simulated clinician. The simulated clinician would use this memory to give diagnostic decisions about specific problem readers.

Procedures in Reading. Given a source for raw data on significant student reading characteristics, three types of diagnoses were considered - those given by the subject teacher using a written decision aid, those generated by simple computerized decision rules, and those given by a complex computerized "clinician."

The diagnoses generated by the reading teacher were generated by a highly structured procedure. The diagnoses were conducted as follows:

1. The subject teacher obtained a copy of the diagnostic and remedial record form used by the Clinical Studies Project of the IRT in their 1980 study (Vinsonhaler, et al, in progress). This record form is organized around the MORAL. It specifically requests a decision on the adequacy of all reading abilties and the information used to

make th

it requ

then pr

2. The sub

key."

factor

and, t

3. In fo

direct

diagno

by exa

the s

elimin

into a

The diagno

result, c

twenty res

this repo

The s

computeriz

diagnostic

threshold.

Performanc

* The time
clinician.
few minute

make the decisions. Then, for each problematic Critical Performance, it requests a list of factors that might have caused the problem and then procedures that can alleviate the problem.

2. The subject teacher took one such record form and created a "scoring key." For this one record, she wrote down her perception of all factors that would cause problems with each Critical Performance and, then, wrote out remedial recommendations (see appendix D).

3. In forming the diagnosis for actual children, the teacher worked directly from the unidentified test battery protocol record. A diagnostic and remedial record form was filled out for the student by examining the student's performance and deciding which subset of the scoring key was relevant for this student. This, obviously, eliminated all problems with translating different student diagnoses into a standardized vocabulary.

The diagnoses performed in this manner were very time consuming.* As a result, only 20 diagnostic reports were written by the teacher. All twenty reports were based on student pre-test performance. Throughout this report, these diagnoses will be called "clinician diagnoses."

The second class of diagnoses were generated by a simple set of computerized decision rules for each of the MORAL Critical Performances. Diagnostic categorization on these factors was determined by a threshold. If the student scored below the threshold, the Critical Performance was declared "problematic;" if above the threshold, the

----

* The time consuming part was the scoring of the test battery by the clinician. Given a scored test battery, the diagnosis would take only a few minutes.

declarat

describe

Critica

Instant
Recogni

Decoded
Recogni

Meaning

Oral Re

Silent

Listeni

In thi

analysi

called

Th

clinici

give d

clinici

MOSES,

of the

declaration was "not problematic." The specific criteria used are described in Table 11.

Table 11
Criteria for Decision Rule Diagnoses

Critical Performance	Measurement Device	Threshold
Instant Word Recognition	SORT Subtest	Number of words recognized instantly = 118
Decoded Word Recognition	SORT Subtest	Number of mediated words = 18
Meaning Vocabulary	Vocabulary subtest	Number of words defined = 10
Oral Reading	Gray subtest	Number of selections with 8 or fewer miscalls = 6
Silent Reading	Durrell or Spache subtest	Number of selections with 40% or more sequential recall = 6
Listening Comprehension	Durrell or Spache subtest	Number of selections with 40% or more sequential recall = 6

In this study the diagnostic decisions were performed by computer analysis of the data. Throughout this report, these diagnoses will be called "rule diagnoses."

The last set of diagnoses generated were done by a computerized clinician designed to collect and analyze data from the test battery and give diagnostic categorizations based on the MORAL. The computerized clinician was run on the Multiple Object Simulated Encounter System (MOSES) developed by the Institute for Research on Teaching and running on the University of Michigan's MTS system. The computer system uses a data base

diagnostic d

elementary i

base. These

problem sc

confirmation

treatment s

The precise

appendix C

diagnoses f

n additio

this same s

The di

investigate

in reading

clinical a

V. Proced

The next r

is, essent

number of

the clini

in treatme

case proc

effectors

must be c

data base of diagnosis, cue, and treatment information to make diagnostic decisions in reading. The computer system has a set of elementary information processing tasks that can be applied to this data base. These elementary tasks are based on the Inquiry model of clinical problem solving and include such elements as hypothesis generation, confirmation and rejection; cue collection; diagnostic termination; treatment selection, and so on (Elstein, Vinsonhaler, & Wagner, 1977). The precise behavior of the computer decision aid is demonstrated in appendix C. The computer decision aid was used to generate pretest diagnoses for the same twenty children diagnosed by the subject teacher. In addition, it was used to generate midtest and posttest diagnoses for this same set of students.

The different procedures for determining the diagnoses were used to investigate the effectiveness and reliability of each diagnostic method in reading. Reliability of diagnostic categorization is central to both clinical and epidemiological research.

IV. Procedures for Treatment Observation

The next requirement for a clinical and epidemiological research study is, essentially, measures of treatment. Treatment may consist of a large number of actions initiated by the clinician and supported by the case, the clinician or others. Therefore, wide variation may, in fact, exist in treatments that are nominally the same. At this point, the model of case process again becomes critical as it highlights the significant effectors of change in case performance. Some measurement instruments must be created for the important effectors of case change.

Gener

the creat

opposed

discrepanc

clinician

the clini

researche

The

variation

the treat

situation

1. It i

2. From
and

3. For
sign
defi

4. Crea
trea
with
incr
effe

5. Use
acti

General Procedure. The purpose of this procedure is to assist with the creation of a treatment observation form to document the actual (as opposed to the prescribed) treatment that a case receives. Any discrepancies between intention and action might be observed by the clinician, the case, or a researcher. In any first attempt to generalize the clinical and epidemiological research paradigms to a content area, researcher observation will be assumed.

The procedure described here can only be suggestive in nature. Wide variation will exist in each field since the concern, now, is observing the treatment received by real cases in a real clinical problem solving situation.

1. It is assumed that a treatment has been selected.

2. From the model of case process, pull out the Critical Performances and their most significant effectors.

3. For each of these case performances and effectors, pull out the most significant effectors of the treatment that will be applied (as defined by the model of case process).

4. Create an observation form of these significant effectors of treatment. A possible organization would be a two dimensional grid with one axis the effectors, and the other time. Then at each time increment, count or otherwise measure the qualities of each effector.

5. Use the form throughout the course of treatment to determine what actions were taken and what the status of significant effectors was.

Proce

to codify

the presc

was create

techniques

the charac

using the

provide a

amount of

time, and

instructor

particular

character

model, li

reading

were the

each stu

terms of

given la

study."

The

the stud

consisted

sections

were lo

task wi

Procedures in Reading. The last measurement device required was one to codify the actual instruction received by the students (as opposed to the prescribed or intended instruction). A treatment observation form was created that was based on the MORAL, research on observational techniques described in the related research section of this thesis, and the characteristics of general remedial reading classroom instruction. Using the Carroll model as a guide, the purpose of the form was to provide a means for recording the types of instructional tasks used, the amount of time spent in each task, the amount of activity during this time, and the corrective and affective feedback provided by the instructors. Because of the limitations imposed on the study by the particular setting, it was felt that this information would adequately characterize the instruction and that the other constructs from the model, like the instructional quality, were essentially constant. The reading clinician assumed initially that all student reading problems were the result of instructional deficits. The instruction planned for each student consisted of the same activities and differed primarily in terms of pace. A complete description of instructional activities is given later in the section entitled "Procedures for Executing the Study."

The treatment observation form evolved during the first month of the study as it was used daily to record instruction characteristics. It consisted of two 8.5x14 inch pages taped together. The pages had sections for each student in the class. Within each student's section were locations for recording how much time was spent on each type of task with what types of feedback. One form was used for each class. The most reliab

materials,

words or pa

on the se

the instruc

four parts

Rarely did

student.

F. Reliabi

report in

the study

As or

the type o

researche

After und

that cod

Evertson,

and the

suffice.

V. Proce

No new p

performs

differen

treatmen

the diag

most reliable information collected included type of task, time on task, materials, and instructor. Less reliable was the estimate of number of words or pages read in a task. Least reliable were measures of feedback on the second page, consisting of type and frequency of feedback from the instructor to the student. Each entry in the form was divided into four parts, one for each of four possible tasks in one class session. Rarely did more than four settings occur in one class session for one student. A sample instruction observation form is included in appendix F. Reliability of classroom observation is described later in this report in the section covering the actual procedures used in executing the study in reading.

As one last check on the adequacy of the coding form for collecting the type of data described, an experienced classroom observational researcher (Linda Anderson) examined the forms and suggested changes. After understanding the data collection needs of this study, she felt that coding schemes that she had used in previous studies (Anderson, Evertson, & Brophy, 1979) were inadequate for the current requirements and the forms that had been created empirically for the study would suffice.

V. Procedures for Outcome Measurement

No new procedures for measurement of outcome are needed. One simply performs a second diagnostic assessment and measures outcome as the difference between performance measures or diagnoses before and after treatment. This is an entirely reasonable definition of outcome since the diagnostic process can distinguish problematic and non-problematic case perfo

V. Proced

Finally, w

planned ar

Gene

paradigms

clinical

epidemiol

the deter

1. the

clin

to d

of i

2. the

step

long

sele

3. the

res

stu

4. the

case performances no matter when the case is seen.

VI. Procedures for Executing the Study

Finally, with all of the groundwork laid, the actual research study was planned and executed.

General Procedure. The clinical and epidemiological research paradigms provide a formal means for the aggregation and integration of clinical experience. The steps necessary to complete a clinical and epidemiological research study (Feinstein, 1977) can be organized around the determination of the following:

1. the objective of the research: the individual clinician, the clinical researcher, or the epidemiological researcher must be able to define what information is desired. This is usually done in terms of initial state, maneuvers, and subsequent state.

2. the "intake" (sampling) of subjects: the series of self selecting steps between an unidentified patient and a research subject may be long. The researcher must specify how the potential subjects will be selected.

3. the maintenance plan: the research study takes place over time. The researcher must have means to keep the selected patients in the study.

4. the initial identification: the general concepts of the research

obje
the
or
meas
decl
5. the
requ
of
Simi
6. the
natu
stud
For
7. the
inte
subj
dete
8. the
dete
the
9. the
shou
10. the
shou

objective must be operationalized into a measurement procedure for the significant characteristics of the case, diagnosis, treatment, or outcome. This can be particularly difficult since many measurements may be the results of subjective, unreliable clinical decision making.

5. the subsequent identification: again a measurement procedure is required. This time it must measure the significant characteristics of the case, diagnosis, treatment or outcome after the maneuver. Similar problems can occur.

6. the allocation of maneuvers: the various maneuvers conducted by nature or man must be allocated to the subjects. For observational studies this consists of determining and classifying what happened. For experimental studies, the maneuver is assigned directly.

7. the handling of intrusions: the maneuver of the study might be interrupted by nature, the experimenter, the clinician, or the subject. The adherance of the study to the desired maneuver must be determined and, potentially, adjusted.

8. the measurement of transition: the outcome of the maneuver must be determined. The transition that occurred from the initial state to the subsequent state must be measured.

9. the analytic procedures: the statistical analysis of the results should be planned.

10. the extrapolation: the final evaluation of the finished experiment should be planned.

underneat
every p
environme
categorie
observe.
case mode

Proc

reading t
readers a
data coll
students'
selected s
received b

Objective

The
research
various
character
personal
treatment
certain
environme
outcome,
of the va

Underneath all of these steps lies the model of the case process. At every point, decisions have to be made about what personal or environmental characteristics of the case to measure, what diagnostic categories to use, what treatments to choose, and what outcomes to observe. The basis upon which all of these decisions are made is that case model.

<u>Procedures in Reading</u>. In this study a junior high school remedial reading teacher was observed while she examined 42 seventh grade problem readers and instructed 30 of these 42 over one school year. The specific data collected included (1) pre-, mid-, and post-test data on the students' reading abilities, (2) teacher and computer diagnoses on selected students, and (3) coded observation of the actual instruction received by the students.

Objective:

The objective of the combination clinical and epidemiological research study in reading was 1) to determine baseline frequencies for various personal, environmental, diagnostic, treatment, and outcome characteristics, 2) to search for associations among diagnosis and personal or environmental characteristics and between outcome and treatment characteristics using retrospective analysis, 3) to test certain hypotheses about the relationship between personal or environmental characteristics and diagnoses and between treatment and outcome, and 4) to detect any problems with the definition or measurement of the various reading characteristics.

intake ar

This

a Lansing

teacher

Collabora

State Un

school,

Institut

of clini

assisted

teaching

Two

subject'

reading

higher,

the rem

students

reading

selected

* A cu

the ent

the sub

----

* Altho
with si
making

Intake and Maintenance of Subjects:

This study was based upon the observation of one reading teacher in a Lansing area junior high school and the 42 seventh grade students this teacher saw during the school year. The reading teacher was a Teacher Collaborator with the Institute for Research on Teaching at Michigan State University and a reading specialist for a Lansing area junior high school. She worked for the Lansing School District each morning and the Institute each afternoon. She was an experienced researcher in the area of clinical decision making (with a Ph.D. in educational psychology) who assisted in the design and procedures of the study as well as the teaching of the children.

Two reading programs were available to problem readers in the subject's school. The first was a skills lab designed to improve the reading of children with an initial reading level of 4th grade or higher. The second, for children with the poorest reading levels, was the remedial reading program conducted by the subject teacher. The students in the study, then, were to be those with the most severe reading deficiences. An initial group of 41 seventh grade students was selected on the basis of Stanford Achievement Test total reading scores. * A cutoff point of 4.0 grade equivalence was used. Those students from the entire 7th grade population who scored 4.0 or lower were referred to the subject teacher. These 41 students were pretested.

----

* Although only 41 students were initially examined, one more student with similar background was referred to the teacher later in the year making the total 42.

Beca

placed

students

1. Two

and

prev

2. Two

thei

3. The

perf

and

comp

inst

con

re-

At

teache

distric

same to

enough

the fun

request

------

* Altho
mean, s
polici

Because of resource constraints, not all 41 students could be placed in the subject teacher's remedial reading program. Thirty students were selected for instruction from the 41 as follows:

1. Two students were rejected because they were repeating seventh grade and had received instruction from the subject teacher in the previous year with little success.

2. Two students were rejected because other programs were available for their special needs - one was bilingual, one was learning disabled.

3. The remaining 37 students were ranked on the basis of pretest performance on the Slosson Oral Reading Test grade equivalence score and the highest graded paragraph attempted on the silent reading comprehension task. The lowest 30 were registered to receive instruction from the subject teacher.* This instruction was to continue at least until midtest when performance measures were to be re-taken.

At the midtest, several students were dropped from the subject teacher's program. One student was dropped because he had moved from the district; one because he was dropped from school; one because he never came to class. Four students were dropped because they performed well enough in class and on the midtest to not warrant further instruction. One further wrinkle concerning tested students was that due to a special request from the subject teacher's building principal, one extra student

----

* Although choosing the lowest 30 would result in regression toward the mean, some defensible criterion had to be used to satisfy external political constraints.

was give

number of

Fina

students

A s

in the fo

student

Descript

Princip
(48)
LD stude
(67)
Bilingu
(4)
Student
(43,14
Student
(5, 1
Student
(9, 1
Student
the re
Totals

The ch

severi

teachin

problem

time.

keep i

theref

was given the pretest battery at this midtest time bringing the total number of students examined to 42.

Finally, before posttest at the end of the school year, three more students were lost due to injury, pregnancy, and chronic absence.

A summary of the students and the data available on each is given in the following table. The numbers in the description column refer to student record numbers.

Table 12
Available Data by Student

Description	Pre	Mid	Post
Prinicipal's student (148)	1	1	1
LD student (107)	1	0	0
Bilingual student (114)	1	0	0
Students previously taught by instructor (143,146)	2	2	2
Students scored too high to teach (115, 125, 127, 130, 137, 145, 147)	7	7	3
Students taught - better by midtest (109, 136, 142, 144)	4	4	4
Students taught the entire year (the rest)	26	25	24
Totals	42	39	34

One characteristic of the students should be emphasized, namely, the severity of their reading deficits. All students examined had extreme reading problems and many had corresponding attitudinal or personal problems. Such problems tend to compound and worsen with the passage of time. Throughout the description of the results of this study, one must keep in mind how poor the reading ability really was, how limited, therefore, were the range of cases, and how the expected outcome in reading

unassiste

initial

The

activitie

1. admi

prob

post-

2. diag

and

3. deve

inte

4. obse

work

5. ana

perf

The pre

1. 9/4,

des

ear

Cri

rel

reading would be a decline relative to grade placement for the unassisted student.

Initial Measures, Subsequent Measures, & Allocation of Maneuvers:

The procedures for executing the study centered around the major activities described earlier.

1. administering the measurement devices to children with reading problems at three points in the year (pre-, mid-, and post-instruction).

2. diagnosing the children's reading problems both by human clinician and by computer.

3. developing and using observational forms to record precise student interaction with actual instruction

4. observing and coding of the subject teacher activities while she worked with her students

5. analyzing the relationships among the various aspects of student performance, student gain, and actual instruction.

The precise chronology of events is as follows:

1. 9/4/80-9/19/80 Clinical Test Battery generation. Based on the MORAL description of the reading process, the test battery described earlier was constructed. It was designed to measure all of the Critical Performances described by the MORAL and as many potentially relevant effecting factors as possible.

2. 9:22

patt

refe

prev

coul

base

reco

sign

clin

data

the

simi

caus

init

of

diff

trea

and

Firs

many

dete

corr

be

3. 10/

ins

2. 9/22/80-10/1/80 Pretesting and Selection of Students. The test battery was administered to 42 seventh grade children who had been referred to the reading teacher on the basis of SAT test scores the previous spring. Due to limited resources, only 30 of the students could be retained and serviced. Placement into the class, then, was based on a small subpart of the pretest, namely the SORT word recognition performance and the comprehension scores. Note that significant limitations were placed on the study here by the clinician. Although very willing to collect the needed diagnostic data about the children, the clinician used only a small subset of the data for diagnostic and treatment decisions. Having taught similar children in previous years, the clinician had a single causal diagnosis that she presumed to be true for all children initially, namely, that their reading difficulties were the result of some previous instructional deficit. The clinician conducted no differential diagnosis; the clinician prescribed the identical treatment for all children, the only difference being starting point and pace.

This restriction on the study can be viewed in one of two ways. First, since the data was collected on all Critical Performances and many effectors, whether or not the clinician was correct could be determined from the empirical data. On the other hand, if she were correct, then this would again limit the range of results that could be gained from the study because the range of students was small.

3. 10/14/80-1/29/81 Instruction and Instruction Observation. The actual instruction took place during this period. Each student was pulled

out

gene

miss

resi

the

stud

day,

grou

the

Micr

stud

adul

twen

coup

stud

stud

leav

with

obse

unde

trea

struc

out of another class for the reading activities. The timing was generally set up so that one-half of a regular school period was missed to receive the special reading instruction. The student was responsible for making up the missed work and the other teachers in the building were aware of the situation and supportive of it. The students were scheduled to come four days a week, for 25 minutes a day. There were between three and six students in each of five groups. The reading teacher had one teaching assistant and often had the additional assistance of a graduate teaching intern from Michigan State University. The instruction ratio was often two students per adult and only rarely went as high as five students per adult.

These students met in a room that measured about twelve feet by twenty feet. Inside this small room were the teacher's desk, a couple of bookcases, chairs and three tables each seating four students, and a couple of chairs for the researcher observers. The students would arrive on the half hour, work for 25 minutes, and leave. The characteristics of this instruction and its interaction with the students were monitored almost every day by a researcher observer.

Actual Instruction:

The treatment provided to the students in this study was not under the control of the research study but, rather, was normal treatment used by the clinician. The instruction was very structured. The clinician made one basic assumption for all of her

stude
in re
being
instr
and
compr
cons
the
char
stud

Sub
seri
firs
basi
one
more
unt
asso
the
con
prov

spe
the

students, namely, that any problems with the Critical Performances in reading were caused by previous instructional deficits. This being the case, the specialist had established a sequence of instructional activities designed to increase a student's instant and decoded word recognition, oral reading, and silent comprehension. The major differences in treatment across students consisted of the point of entry into the instructional sequence and the pace. Note that such a restricted range of treatment characteristics placed a severe limitation on the outcome of the study itself.

The sequence of instructiona activities was centered around the Sulliven reading series (Sullivan Comprehension Readers, 1980). This series consists of short readers setup in twenty major levels. The first level is written with words requiring the use of only the most basic and clear phonetic regularities. The second level introduces one or two more phonetic redundancies. Each level introduces even more of the phonetic system, at progressively more difficult levels, until the twentieth level contains most of the sound-symbol association rules of the language. For students that progressed past the Sullivan series, grade level basal texts were available for continued reading practice. These readers and basal texts, then, provided the contextual reading material for instruction.

From each source of contextual reading material, the reading specialist drew up lists of words that characterized the text. In the Sullivan series the chosen words emphasized the new phonetic rules being introduced. Words from basal texts also reflected any new phonetic regularities as well as other word mediation strategies such as use of affixes. These word lists served as the materials for isolated word practice.

The isolated word lists and the contextual reading material were used by the reading specialist in both silent (or group) activities and oral (or individual) acitivies. The major activities of the class consisted of the following:

a. Oral reading of isolated words: A group of students were all given a copy of a typed word list. The teacher would then go in round-robin style asking each student in the group to read in succession a certain number of words. The round-robin would continue until all the words were read. For example, the students might be working with a four-page list of 200 words. The instructor would ask the first student to read the first five words aloud. While the first student read these five words, the other students were variously instructed to listen or look ahead to their words. When the first student had completed the five words, the second student would be asked to read the next five. Then the third student would read the next five and so on.

   As the students tried to read each word, the instructor would provide feedback on the correctness of the attempt. Correct performance was often signaled with no feedback or some one word positive comment. Incorrect performance was indicated by feedback attempting to focus the student's attention, recall a known decoding rule, etc.

b. Silent reading of isolated words: At times when class management constraints would not allow the individual attention required for oral word list use, a student or set of students would be asked to read a given word list silently and mark all the words that "gave them trouble." When the students were done, they were to inform the instructor. At this point the instructor would assist the students with all marked words as well as spotchecking some of the unmarked words.

c. Oral reading of contextual materials: A group of students were all given the same or occasionally different contextual material. The materials generally consisted of stories with 1-4 paragraphs per page, but this varied widely over materials of different difficulty levels. The teacher would again go round-robin through the group of students asking each one to read 1-2 pages of text at a time. The reading would continue or stop at the teacher's discretion. She had some criteria for determining adequace of oral reading performance and would try to keep the group reading consecutive pages until it was reached. While one student was reading the other students were variously asked to follow along or to silently read ahead the pages for their turn.

As the students read the selections aloud, the instructor provided corrective feedback on a variety of reading performances. The major concerns were correctness of word recognition, inflection, fluency, and speed/ Incorrect performance was generally indicated by global comments and cues to help the student correct the performance unaided.

d. Silent reading of contextual materials: Students singly or in a group were asked to read through selections of connected text. When the students were done they were to inform the teacher. Often students were asked to complete a comprehension form that required the students to identify the main characters, the setting, the sequence of events, the major problem in the story and its resolution.

e. Group instruction: Only rarely did the teacher attempt to provide direct instruction to the group. At such times the topics tended to be directed at word recognition.

f. Homework: During the first part of the year, homework was assigned and monitored. However, as the term progressed, it became apparent that few of the students completed such assignments and their importance faded.

Within the materials and activities just described, the instructor used only one sequence of instruction, specifically, practice of the word list until mastery was demonstrated orally and, then, practice of the associated connected text until mastery was demonstrated orally. When the materials at one level were complete the next level would begin. Oral and silent acitivies varied with the number of people in class and the number of assistants that the clinician might have; oral reading seemed to be preferred, silent reading was the sometimes necessary alternative. In general the instructor would try to keep the students in one class on the same materials.

As can be seen, across all instructed students the quality of instruction and opportunity for instruction was essentially the same. During classroom observation the ability of the students to understand the instruction did not seem to differ greatly. The two major differences seemed to be in the students' aptitudes for reading and their perseverance in the instructional tasks.

Observation

The observation of classroom treatment was conducted by one of two observers seated in the classroom in full view of the students. The observation was conducted primarily during the first part of the year from pre-test time to mid-test time.

The observer sat at a desk in one corner of the rectangular room. From this corner the observer watched three tables of students. One table was directly ahead of the observer, one directly to the right and one diagonally in between. Each table was within five feet of the observer.

The observers were introduced to the class at their first observational session but were from then on essentially ignored. Observers never spoke with the students on a typical day. As far as could be determined, the existence of an observer did not seem to change student behavior. For example, if a particular student was supposed to be working silently but was instead staring off into space, this time wasting behavior rarely changed if the student saw the observer watching.

Many different observation forms were tested before the final one given in appendix F was completed. As the instruction progressed it was apparent that the amount of time provided for each student's instruction, as well as the quality of feedback were essentially constant across students. Aptitude for instruction was not under control of the instructor. Hence, the major instructional determiner seemed to be under the control of the student, namely, perseverance. The form used to collect the data reflects this perspective keying in primarily on engaged time. As the obervation continued, it became clear that, again, the range of activities for the students was quite structured and essentially constant. To make sure that nothing critical was missing from the observation device, an experienced classroom researcher (Linda Anderson) examined the forms, suggested some changes, and stated that for the types of data desired here, the coding schemes that she had used in the past were inappropriate (Anderson, Evertson, and Brophy, 1979).

Attempts were made to estimate the reliability of the coders by having both observers collect data on the same classroom session. Because of the extremely cramped quarters available, it was discovered that the two observers themselves interfered with each other. One observer could not help but notice as the other observer redirected attention from one group to another; the reliability measures obtained, therefore, are undoubtedly inflated. However, looking at the low inference nature of the tasks being recorded and the systematic way in which it was performed, possibly the correlation among observed events of 0.85 was not too far from the correct amounts.

4. 2/2/81-2/13/81 Midtesting. The test battery was administered again to all available children. Three students had been lost. Alternate forms were used for all subtests.

5. 2/16/81-5/19/81 Instruction and Instruction Observation. Instruction proceeded essentially as before. Observation was greatly reduced. Some students were dropped from further participation because their reading had improved substantially.

6. 5/20/81-6/5/81 Posttesting. The final test battery was administered to all children again.

In addition to this schedule of events, at times throughout the year, the following took place:

1. Written diagnosis of one-half of the pretest performances by the reading specialist

2. Computer diagnosis of pre and post test performances

Summ

The pr

aspec

parad

study

centr

spec

thes

deci

trea

deci

let

Summary

The procedures described above detailed some of the more important aspects of generalizing the clinical and epidemiological research paradigms to any field as well as the specific procedures used in this study of reading diagnosis and treatment. More than anything else, the central concern of these procedures is the long-term aggregation of specific, reliable, model-directed measures of clinical experience. As these descriptors of empirical experience accumulate, a data base for decision making is formed. The data base can assist in the diagnostic, treatment, and follow-up decisions of a clinician or in the policy decisions of an administrator by providing access to past experience. Let us turn, now, to the beginning of such a data base in reading.

Introdu

The res

clinic

clinic

in rea

second

for i

the r

data

all

child

child

reco

two

the

lea

dis

be

CHAPTER 4

Results

Introduction

The results of achieving the first objective, that of determining how a clinical and epidemiological research study could be completed in any clinical problem solving field in general and carrying out such a study in reading, were described in Chapter Three. This chapter centers on the second objective, that of evaluating the results of the reading study for its comments on the field of reading and on attempting to generalize the research paradigms outside of medicine.

The data that were available for analysis included raw performance data on the pre-, mid-, and post-tests, the decision-rule diagnoses for all children on all tests, the subject teacher's diagnosis of 20 children's pre-test performance, the computer's diagnosis of the same 20 children on the pre-, mid-, and post-tests, and the coded observation record of actual instruction. Before the data were analyzed, however, two checks had to be made.

The first area of concern was the equivalence of different forms of the reading subtests in the battery. Because of the possibility that learning could occur from pre- to mid- to post-test, various forms of different stimulus materials were prepared. The first analysis that was performed, then, was to test the assumption that the different forms of

each subte
indicated
the measur
were subst
raw score
comparable
needed tra

Once
adjusted,
of the c
research
environme
reliabil
apparent
the reli
elaborat

For
epidemic

Epidemi

Th
model's
past cl
any re
in the

each subtest were, in fact, equivalent. The results of this analysis indicated that all word recognition measures were equivalent but that the measures of silent reading comprehension and listening comprehension were substantially different. Therefore, adjustments were made to the raw scores for silent reading and listening to make the results more comparable. A complete description of the equivalence of forms and needed transformations is given at the end of this chapter.

Once the equivalence of alternate forms was established and adjusted, the second major concern was the reliability of the data. One of the critical requirements for the clinical and epidemiological research paradigms is reliability in the determination of personal, environmental, diagnostic, treatment, and outcome characteristics. The reliability of diagnostic categorization of Critical Performances can apparently be quite high because the rule diagnosis is limited only by the reliability of the subtests themselves. The reliability concerns are elaborated at the end of this chapter.

For now, let us turn to the analysis of the clinical and epidemiological data in reading.

Epidemiological Data Analysis: P,E -> D

The clinical and epidemiological research paradigms are formal models for improving future clinical decisions based on experience with past clinical decisions. It is altogether appropriate, then, to organize any research guided by these paradigms around the major decisions made in the clinical setting.

T

They

those

reta

selec

stude

decis

child

trea

rese

of t

cli

dia

nat

and

ep

The specific decisions that a reading specialist makes are many. They include choosing which children to examine for problems, diagnosing those children, selecting from the diagnosed children the ones to be retained for instruction, grouping the children for instruction, selecting and administering and adapting treatment, determining when a student is "cured" or "beyond help," and so on. However, all of these decisions center around (1) the ability to determine the state of a child's reading and (2) the determination of the most appropriate treatment plan for a problem reader. They focus on the epidemiological research concerns of relating personal and environmental characteristics of the student to a diagnostic categorization (P,E -> D) and around the clinical research concerns of relating characteristics of the student, diagnosis, and treatment to outcomes (P,E -> D -> T -> O). Taking the natural chronological order, a clinician determines the diagnosis first and then chooses a treatment. Hence, let us begin with the epidemiological analysis of diagnostic categorization.

<u>Characteristics of Diagnosis and Cross-sectional Analysis</u>. The most straightforward analysis of epidemiological data is simply counting without reference to cause or effect, a so-called "cross-sectional" analysis. This type of analysis provides information on the prevalence of certain data in the populations under study. In this study, the sample consisted of all 7th grade students in a junior high school who had received a grade equivalent score in reading of 4.0 or lower on the Stanford Achievement Test, a group of very poor readers. The most important prevalence rates are those of diagnostic categories

and dia

Th

availab

compute

made i

diagno

Perfor

clinic

Perfor

E

were

less

of o

were

above

prob

be e

fact

the

achi

dia

giv

and diagnostic profiles.

Three different sources of diagnostic categorization were available: the clinician, the computer decision rules and the computerized clinician. The clinician and computerized clinician both made initial diagnoses for 20 of the 42 children. The clinician's diagnosis included categorizations on all six of the MORAL Critical Performances as well as 30 other causal factors. The computerized clinician's diagnosis also included decisions on the six MORAL Critical Performances and on 58 other causal factors.

Because the diagnostic classifications by clinician and computer were not available for all cases, and because a classification using less inference was desired, a third diagnosis was prepared using a set of decision rules based directly on raw performance data. Thresholds were established for the Critical Performances. If a student scored above the threshold, the diagnostic categorization would be "not problematic;" below the threshold, "problematic." Such thresholds could be established only for the Critical Performances and not for any causal factors. This rule-bassd diagnosis and the raw performance data will be the major dependent variables throughout the rest of the analysis.

A comparison of the prevalence of Critical Performance problems across diagnostic sources is given in Table 13. The percentage of diagnosed cases with a problem in a particular Critical Performance is given along with the exact ratio of problematic cases to total cases.

Instant Wo
Recognitio
Problem

Decoded Wo
Recogniti
Problem

Meaning
Vocabular

Oral Read

Silent Re
Comprehe

Listenin
Comprehe

Wi

appears

trouble

diagnos

actual

Table 13
Prevalence of Critical Performance Problems

	Clinician Diagnosis (N=20)	Rule Diagnosis (N=43)	Computerized Clinician Diagnosis (N=20)
Instant Word Recognition Problem	95% (19/20)	81% (34/42)	60% (12/20) (No decision possible on 6 cases)
Decoded Word Recognition Problem	100% (20/20)	90% (38/42)	85% (17/20) (No decision possible on 3 cases)
Meaning Vocabulary	75% (15/20)	67% (28/42)	70% (14/20) (No decision possible on 6 cases)
Oral Reading	95% (19/20)	76% (32/42)	95% (19/20)
Silent Reading Comprehension	100% (20/20)	100% (42/42)	50% (10/20) (Differing definition of diagnosis)
Listening Comprehension	75% (15/20)	92% (39/42)	25% (05/20) (inappropriate weighting of diagnostic information)

Without even attempting to correlate the different diagnoses, it appears that students at this level of SAT performance generally have trouble with all Critical Performances in reading. The different diagnostic sources generally agree on this. The major differences in actual percentages seem to be due merely to the computerized clinician's inability

However, t

the diffe

listening

computeri

clinician

the algo

can be at

silent r

comprehen

similar

placemen

"reading

second p

reading

compute

clinici

Wi

expect

be high

propor

possib

descri

-----

* The

was to

this.

inability to categorize some cases on some Critical Performances.* However, there are a few significantly different prevalence rates for the different diagnostic sources. Note especially silent reading and listening comprehension. The prevalence rates for the diagnosis by the computerized clinician is one-half and one-fourth of that for the clinician or rule diagnoses, respectively. After careful examination of the algorithm used by the computerized clinician, this large difference can be attributed to a disagreement on the definition of problematic silent reading and inappropriate weighting of data for listening comprehension. The clinician and the decision rule evidently used a similar criterion - something like "reading more than 1 year below grade placement." The computerized clinician algorithm used the definition "reading more than one year below word recognition level." From this second perspective, there were a number of students (10) whose silent reading comprehension was limited by their ability to read words. The computerized clinician algorithm did not call this problematic; the clinician and rule diagnoses did.

With these minor differences in high prevalence rates, one would expect the measures of association among the three diagnostic sources to be high. The measure used to compare the relative diagnoses is a simple proportion - the number of actual agreements divided by the number of possible agreements. It is based on the simple two by two table described in Table 14.

----

* The computerized clinician had the option to decide that the evidence was too mixed to make a decision. The rule based decision did not allow this. The human clinician never took this option.

Source
Catego

Source
Catego
Proble

The s

comput

cases.

T

report

Instan
Recogn

Decoded
Recogni

Meaning
Vocabul

Oral Re

Silent R
Comprehe

Listenin
Comprehe

Table 14
Agreement Among Diagnostic Sources: The Statistic Definition

	Source I: Category M Problematic	Source I: Category M Not Problematic
Source J: Category M Problematic	a	b
Source J: Category M Not Problematic	c	d

The statistic is computed as (a+d)/(a+b+c+d). Note that this statistic computes agreement on one particular diagnostic classification across cases.

The proportion data for the Critical Performance diagnoses is reported in Table 15.

Table 15
Agreement Among Diagnostic Sources: Critical Performances

	Clinician and Computerized Clinician agreement	Clinician and Decision Rule agreement	Computerized Clinician and Decision Rule agreement
Instant Word Recognition	65% (13/20)	95% (19/20)	70% (14/20)
Decoded Word Recognition	85% (17/20)	95% (19/20)	90% (18/20)
Meaning Vocabulary	75% (15/20)	80% (16/20)	75% (15/20)
Oral Reading	90% (18/20)	85% (17/20)	85% (17/20)
Silent Reading Comprehension	50% (10/20)	100% (20/20)	50% (10/20)
Listening Comprehension	50% (10/20)	80% (16/20)	30% (6/20)

A giv

betwe

diagn

insta

Note

the r

inabi

diffe

For a

judge

Criti

colum

3) ins

vocab

6) lis

corre

probl

A given proportion, like 65% for instant word recognition agreement between clinician and computerized clinician, means that of all students diagnosed by both clinician and computer, agreement on the status of instant word recognition performance occurred for 65% of the students. Note the very close correspondence between the clinician's diagnosis and the rule diagnosis. Clearly, because of the computerized clinician's inability to categorize certain students, its diagnoses are the most different of the three.

Of course, these diagnostic categories do not occur in isolation for a student. Instead the clinical decision maker gives simultaneous judgements on many classifications for one student. The profiles of Critical Performances are defined in the following table. The first column gives the profile - the six symbols indicate the status of 1) instant word recognition, 2) decoded word recognition, 3) meaning vocabulary, 4) oral reading, 5) silent reading comprehension, and 6) listening comprehension in order. A "P" indicates that the corresponding Critical Performance is problematic; an "N" indicates not problematic.

Diagnos

1. P P
2. P P
3. P P
4. P P
5. N P
6. P P
7. P P
8. N P
9. N P
10. N P
11. N P
12. N N
13. N N
14. N N
15. P P
16. P P
17. N P
18. P N
19. N P
20. P N
21. N P
22. P P
23. N P
24. N N
25. P P

T

to und

skills

Profi

skills

Mediat

for t

Descrip

Table 16
Agreement Among Diagnostic Sources: Profiles

Diagnostic Profile	Number of Clinician Cases with Profile (N=20)	Number of Rule Cases with Profile (N=43)	Number of Computerized Clinician Cases with Profile (N=20)
1. P P P P P P	60%	55%	0%
2. P P P P P N	10%	0%	25%
3. P P N P P P	10%	14%	5%
4. P P N P P N	10%	0%	10%
5. N P N P P P	5%	0%	0%
6. P P P N P N	5%	0%	0%
7. P P P P N P	0%	0%	10%
8. N P P P N N	0%	0%	10%
9. N P P P N P	0%	0%	5%
10.N P P P P P	0%	0%	5%
11.N P N P N N	0%	0%	5%
12.N N P P N N	0%	0%	5%
13.N N N P P N	0%	0%	5%
14.N N P N N N	0%	0%	5%
15.P P P P N N	0%	0%	5%
16.P P N P N N	0%	0%	5%
17.N P N N P P	0%	7%	0%
18.P N N P P P	0%	5%	0%
19.N P P N P P	0%	5%	0%
20.P N P N P P	0%	2%	0%
21.N P N N P N	0%	2%	0%
22.P P N N P P	0%	2%	0%
23.N P P P P N	0%	2%	0%
24.N N N N P N	0%	2%	0%
25.P P P N P P	0%	2%	0%

The diagnostic profiles used by the clinician are relatively simple to understand. Eighteen of 20 students have problems with all reading skills and vary only on their language experience potential for reading (profiles 1 - 4). Profile 5 is a student whose instant word recognition skills and meaning vocabulary are intact but who has a problem with mediating strategies for word recognition that might well be responsible for the problems with connected text. Profile 6 seems less likely. It describes a student who has problems with all forms of word recognition and comp

The

indicate

profiles

recallin

problem

decisio

of 20.

detecte

problem

profil

are to

diagnos

differe

Th

by de

these

clinic

the co

T

profil

catego

compar

diagno

earlie

and comprehension yet manages to read aloud adequately.

The profiles generated by the computerized clinician seem to indicate a greater diversity of student ability. Only eight of its profiles overlap at all with the clinician's diagnoses. However, recalling the different definitions for a silent reading comprehension problem, and the cases for which the computer was unable to make a decision, the overlap could probably have reached as high as 12-14 out of 20. Of particular note, however, is that the computer diagnosis detected (supposedly) many students with no instant word recognition problems (profiles 8 - 14) and one student with really minimal problems (profile 14). Knowing that the computer algorithm and the decision rules are totally reliable (i.e., given the same data twice, the same diagnosis will result), these data are again demonstrating apparently different definitions of when a problem exists.

The profiles of the diagnoses generated directly from the raw data by decision rules agreed most closely with the clinician. Almost 70% of these diagnostic profiles were accounted for by those first four clinician profiles. The rule-based diagnoses overlapped very little with the computer but did generate nine profiles not seen elsewhere.

The data just described give the prevalence rates of various profiles. However, it did not directly compare total diagnostic categorization on the same case across diagnostic sources. When comparing the diagnosis of a specific student generated by two diagnostic sources, a single measure similar to the proportion described earlier can be computed. This statistic is also a proportion. But inst

cate

agre

the

ave

on

acro

in p

Ori
Per
Dia

Tot
of
Per
Cau

Ful

the

the

resp

Pag

instead of being a proportion of agreement on one diagnostic categorization across cases, this new statistic is the proportion of agreement on one case across diagnostic categories. This proportion can then be averaged across all students diagnosed by the two sources. This average proportion describes how well the two diagnostic sources agree on total diagnoses across students. This measure of total agreement across cases is contained in Table 17. Standard deviations are included in parentheses.

Table 17
Average Agreement on Total Diagnosis Across Cases

	Clinician and Computerized Clinician Diagnosis	Clinician and Rule Diagnosis	Computerized Clinician and Rule Diagnosis
Critical Performances Diagnosis	0.69 (0.22)	0.89 (0.11)	0.67 (0.17)
Total Diagnoses of Critical Performances and Causal Factors	0.34 (0.06)	not available	not available

The diagnoses of Critical Performances by clinician and decision rule are very close, agreeing 89% of the time. It is not surprising, then, that the agreements of the computerized clinician's diagnosis with the clinician and rule diagnoses were essentially the same, 69% and 67% respectively.

The second row of the table deals with the more fine-grained diagnosis by computerized clinician and clinician only. The computer used a total of 65 diagnostic categories; the clinician used 36. These two di

reading

only

clinic

catego

diagnos

elimina

67% (s

diagno

Da

There

Perform

differ

seemed

differ

criter

perfor

was pr

interv

a deci

result

that t

data

to mat

Or sh

diagnos

two differing vocabularies for describing diagnostic categories in reading were combined into 43 categories. Of these 43, 11 were mentioned only by the computerized clinician, 10 were mentioned only by the clinician. Therefore, there was automatic disagreement on 21 of the 43 categories. This is the reason for the low agreement on the total diagnosis. If the 21 categories on which agreement was not possible were eliminated, then, the agreement between clinician and computer reached 67% (s.d.=11%), approximately equal to that on the Critical Performance diagnosis alone.

During this first analysis, one fact became exceedingly clear. There is no well defined, operational construct for any of the Critical Performances of the model or any of their effectors. The major differences among diagnostic categorization from the three sources seemed to come from differences in definition of the construct and differences in levels of inference. Apparently, the clinician used criteria very similar to the decision rule, namely, comparing performance on one subtest to some threshold. The computerized clinician was programmed to attempt more subtle levels of diagnosis by combining interview and selected performance data across reading tasks when making a decision. The two different diagnostic approaches give different results. It can be argued that the most reasonable interpretation is that the computerized clinician system has never been "fine-tuned" on data from real children. The weighting system it uses should be revised to match more closely the clinician's diagnosis or the rule diagnosis. Or should it? Indeed, looking at the raw data, the clinician and rule diagnoses appear more reasonable. But this discrepancy has made it abunda

which

resear

resear

defini

a co

hypot

for c

cause

means

categ

facto

causa

of

in t

these

natu

measu

(by

only)

first

abundantly clear that reading has no "pathologist" or "lab test" to tell which is right!

It seems apparent that for the clinical and epidemiological research paradigms to succeed for any individual clinician or large research organization, to as great an extent as possible, the definitions of all constructs must be operationalized and maintained as a constant over time. For each construct, one must define the hypothesized function, specific procedures for measurement, and methods for combining measures into a final judgement.

Retrospective Analysis. The next epidemiological analysis examines cause and effect by looking backwards - a retrospective analysis. This means that groupings must be formed on the basis of diagnostic categorization and followed to differences in personal and environmental factors (D -> P,E). This type of analysis can assist in the search for causal hypotheses.

The diagnostic categories examined were the Critical Performances of the MORAL model of reading. Although more categories were available in the clinician and computerized clinician diagnoses of some students, these more fine-grained, causal classifications are more inferential in nature. Since these classifications were not the direct target for measurment, many of the decisions were made on what few data there were (by clinician and computer) or by prior likelihoods (the clinician only). Therefore, these secondary classifications were ignored for this first analysis.

S

the s

consid

clinic

concer

words

fluent

percep

Table

(which

Diagno
Decisi

Instan
Recogn
C
E

Decode
Recogn
C
E

Oral p
C
E

Silent
Compre
C
E

T

of stu

Some of the personal and environmental characteristics available on the students included interview data, sex, and SAT scores. First consider the students' self perception of their reading ability based on clinicial interview. During interview, the students were queried concerning their self perception of their ability to 1) recognize the words in their school books instantly, 2) sound out words, 3) read fluently, and 4) understand what they read. The association of these self perceptions with the student's diagnostic categorization is presented in Table 18. The diagnostic source used was the decision rule diagnosis (which very closely paralleled the clinician's diagnoses).

Table 18
Differences in Self Perception Based on Diagnosis

Diagnosis Based on Decision Rules	Self Perception Good	Self Perception Bad
Instant Word Recognition		
Good	50% (1 of 2)	50% (1 of 2)
Bad	56% (10 of 18)	44% (8 of 18)
Decoded Word Recognition		
Good	0% (0 of 1)	100% (1 of 1)
Bad	21% (4 of 19)	79% (15 of 19)
Oral Reading		
Good	0% (0 of 4)	100% (4 of 4)
Bad	19% (3 of 16)	81% (13 of 16)
Silent Reading Comprehension		
Good	--	--
Bad	35% (7 of 20)	65% (13 of 20)

The most obvious characteristics of these data is that the number of students without a problem for any given diagnostic category is rea

the

prob

that

stud

This

thos

who

part

unre

inte

stud

stud

thou

self

good

sex,

use

grou

them

Basi

vari

really too small in this study to give a meaningful comparison. None of the differences among self perception for students with and without problems were at all statistically significant, but with the small N that is not surprising. The general trend was that the majority of students had a poor self perception across the Critical Performances. This was true both for students who in fact had problems as well as for those who performed adequately. What is likely is that these students, who are each having .problems in some reading area, generalize their particular weakness to the entire reading act. This is not entirely unreasonable given the way in which the various Critical Performances interact with each other. The only exception is the high confidence poor students have in in their ability to recognize words instantly.

In addition, for every case there was a greater percentage of poor students who thought they were good than there were good students who thought they were poor. In fact, the data indicate that the student's self perception (measured as it was through interview) is not at all a good indicator of correct diagnostic categorization.

One other check was run to see if another demographic variable, sex, had confounded the results. Clinical and epidemiological research use a simple technique called "standardization" to compensate for groupings that may differ on some significant personal or environmental characteristics other than those under study (Roberts, 1977, p.148). Basically, standardization is performed by stratifying the analysis on various levels of the confounding characteristic.

F

categ

Diagno
Male S

Instan
G
B

Decode
G
B

Oral R
G
B

Silent
G
B

Diagno
Female

Instan
G
B

Decode
G
B

Oral R
G
B

Silent
Go
Ba

Reexamining the association of self perception and diagnostic categorization in light of sex differences gives the data in Table 19.

Table 19
Differences in Self Perception Based on Diagnosis by Sex

Diagnosis Based on Decision Rules for Male Students	Self Perception Good	Self Perception Bad
Instant Word Recognition		
Good	1	1
Bad	9	3
Decoded Word Recognition		
Good	0	1
Bad	4	9
Oral Reading		
Good	0	4
Bad	3	7
Silent Reading Comprehension		
Good	0	0
Bad	4	10

Diagnosis based on Decision Rules for Female Students	Self Perception Good	Self Perception Bad
Instant Word Recognition		
Good	0	0
Bad	1	5
Decoded Word Recognition		
Good	0	0
Bad	0	6
Oral Reading		
Good	0	0
Bad	0	6
Silent Reading Comprehension		
Good	0	0
Bad	3	3

Th

st

to

ca

mu

ch

ca

es

ab

Pe

di

The

im

ove

The

mea

or

det

for

Des

abi

This stratification indicates that in this sample of students the female students had a consistent (and correct) poor self assessment. It appears to be the male students that are incorrectly assessing their ability.

Prospective Analysis. The final epidemiological analysis examines cause and effect by looking forward, a prospective analysis. Groupings must be formed on the basis of personal and environmental characteristics and followed, then, to differences in diagnostic categorization (P,E -> D). This type of analysis can be used to directly estimate the risk of contracting a certain problem given the presence or absence of a particular personal or environmental factor.

The diagnostic categories examined were the MORAL Critical Performances again. The major personal factor used to predict the diagnostic categorization were two estimates of language experience. These two estimates were based on the judgement of test administrators immediately after test completion. They were estimates of a student's overall language ability and of the student's mean length of utterance. The overall language ability was judged to be high, average, or low; the mean length of utterance was judged to be a word, a phrase, a sentence or multiple sentences. The MORAL views language experience as a major determiner of reading ability and this analysis tested the association for these students.

The relative diagnostic categorization of the groups defined by the test administrator's judgement of high, average, and low language ability is contained in Table 20.

Table 20
Differences in Rate of Problem Diagnosis Based on Ratings of Language Ability

	High Language Ability (N=4)	Average Language Ability (N=17)	Low Language Ability (N=21)
Instant Word Recognition	50.0%	70.6%	95.2%
Decoded Word Recognition	50.0%	94.2%	95.2%
Meaning Vocabulary	0.0%	58.5%	85.7%
Oral Reading	50.0%	64.7%	90.5%
Silent Reading	100.0%	100.0%	100.0%
Listening Comprehension	75.0%	88.2%	100.0%

The most obvious aspect of these data is the increasing rate of problems with all Critical Performances as the test administrator judges language to be increasingly deficient. Before discussing these data further the data on the second judgement of language ability, that of mean length of utterance, is presented in Table 21. In this table "Multiple" refers to the rating of "Mean Length of Utterance is Multiple Sentences." Similarly, "Single" refers to "Mean Length of Utterance is Single Sentence," "Phrase" refers to "Mean Length of Utterance is Phrase," and "Word" refers to "Mean Length of Utterance is Word."

Table 21
Differences in Rate of Problem Diagnosis Based on Ratings of Mean Length of Utterance

	Multiple (N=9)	Single (N=8)	Phrase (N=13)	Word (N=12)
Instant Word Recognition	55.6%	75.0%	84.6%	100.0%
Decoded Word Recognition	88.9%	87.5%	84.6%	100.0%
Meaning Vocabulary	55.6%	62.5%	61.5%	83.3%
Oral Reading	66.7%	75%	61.5%	100.0%
Silent Reading	100.0%	100.0%	100.0%	100.0%
Listening Comprehension	88.9%	87.5%	92.3%	100.0%

This second table generally demonstrates the same trend as the first. Apparently, language ability is a large determiner of diagnostic categorization. The differences between consecutive groups is not statistically significant but differences between the highest and lowest judgements are statistically significant at the 0.05 level for instant word recognition and meaning vocabulary.

<u>Summary</u>. Throughout this epidemiological analysis, several issues are raised. The first and simplest is a matter of sample size. For any of these analyses to provide guidance in the accurate diagnosis of reading problems or the eventual prevention of reading problems, quantities of data on non-problem readers must be collected. The relative specificity and sensitivity of test for diagnoses cannot be determined without such data.

pri

wit

ass

not

stu

mig

env

stu

ex

the

ans

for

rea

di

on

stu

ser

pro

sta

thro

read

if (

yea

There are two more serious problems that the data reported here bring out. The first issue is causality. The data may, indeed, assist with the accurate diagnosis of reading problems by providing empirical associations among the various data. However, the findings actually did not examine causality. To study causality, one must in some way examine students before they develop the reading problems. Prospective studies might group children without reading problems on some personal and environmental characteristics and follow them forward until some students developed problems. A retrospective study would look back from existing problems to some point before the problem arose. This requires the examination of a store of pre-existing records.

Based on the experiences connected with conducting this research, answering questions of causality may be more difficult for reading than for medicine. Just one obvious complication when looking for causes of reading problems is that there is a large cadre of reading teachers with differing abilities, philosophies, instructional styles, and so on. Not only do different students encounter very different teachers but one student can, through the course of a school career, be remediated by a series of teachers holding widely divergent views of what constitutes a problem and its proper treatment. In some sense, there is no natural state. Conscious attempts to alter reading ability begin and continue throughout school. How, then, can all of these effects be disentangled?

The last question raised by these data is the separability of reading problems. First, consider separation as to severity of problem. If one seventh grade student has an instant word recognition level one year below placement and another seventh grade student has an instant word recognition level two years below placement, do they have the same problem? How about levels of one year and four years? Levels of one year and one-half year? If a second grade student is one-half year behind and a third grade student is one and one-half years behind, do they have the same problem?

Next, consider separation based on Critical Performances. The MORAL says that a prerequisite for silent reading comprehension is word recognition. Is it appropriate, then, to diagnose as the computer system did and say that if silent reading performance matches word recognition performance then no comprehension problem exists? Or is the clinician's approach more valid?

Hopefully, this discussion has raised some of the more pressing problems introduced by inadequate definition of constructs and measurements.

Clinical Data Analysis: D -> T -> O

The epidemiological analysis examined the first part of the clinical encounter dealing with diagnostic categorization based on personal and environmental characteristics. The remaining part to be examined is the clinical analysis of outcomes of treatments for particular diagnosis.

Characteristics of Treatments and Outcomes. Although a cross-sectional analysis of clinical data is not possible (because treatment and outcome of treatment cannot exist at the same time), desc

disc

firs

rec

Thi

stu

of

dee

inc

the

sec

gro

sch

stu

Gro

Gro

Gro

Gro

Gro

descriptions of the ranges of treatment and outcomes must surely precede discussion of the retrospective or prospective data analysis.

There were nominally three treatment groups in this study. The first group included twelve students who for various reasons were to receive no treatment with the reading clinician at all during the year. This group will be called "Group None." The second and third group of students, 30 in number, started out together. However, by mid-year four of these students had progressed well enough that they were no longer deemed in need of assistance. Therefore, the second treatment group included four students who were instructed by the reading clinician for the first half of the year but received no such training during the second half. This second group will be called "Group Half." The third group included 26 students who were instructed throughout the entire school year. This last group will be called "Group Full."

The easiest indicator of the amount of instruction that each student in each group received is the number of days in attendance. Group None did not attend at all. Attendance patterns for Group Half and Group Full are in Table 22.

Table 22
Mean Days of Instruction by Treatment Group

	First Half of Year	Second Half of Year
Group Half	42.5 (s.d. 7.0) (range 33-49)	0
Group Full	40.3 (s.d. 7.8) (range 18-49)	24.6 (s.d. 12.7) (range 0-39)

Group None - no instruction at all

dav

info

spen

sile

hump

also

prog

Full

each

During the first half of the year, detailed records were kept each day on the precise instruction received by the children. Such information as the total amount of time in class, the amount of time spent on word lists orally and silently, reading connected text silently, etc., was recorded. Although not as reliable, estimates of numbers of words practiced in isolation and pages read in context were also recorded. Again, Group None received no instruction in the reading program. The characteristics of instruction for Group Half and Group Full are contained in Table 23. The range of actual treatments within each group of instruction is large.

Tota
in m

Time
Ora

Time
Sile

Time
Read

Time
Read

Time
Summ

Cost

Numb
(sc

Numb

Numb

Grou

This

Minu

Reca

there

spen

recog

Table 23
Characteristics of Instructional Time by Treatment Group

	Group Half	Group Full
Total Time in minutes	801 (194) (range 542-1000)	757 (177) (range 230-1053)
Time on Isolated Words Orally	161 (31) (range 127-201)	220 (78) (range 49-405)
Time on Isolated Words Silently	10.3 (12.7) (range 0-26)	24.4 (27.8) (range 0-117)
Time on Contextually Reading Orally	125 (77) (range 39-221)	205 (81) (range 25-380)
Time on Contextual Reading Silently	169 (72) (range 65-223)	87 (25) (range 27-136)
Time on Comprehension Summaries	77 (90) (range 20-211)	24 (34) (range 0-125)
Lost Time	13.5 (13.9) (range 1-30)	9.4 (8.5) (range 0-28)
Number of Words in Isolation	2609 (429) (range 1088-2134)	1511 (534) (range 553-2699)
Number of Pages Orally	141 (71) (range 69-238)	172 (86) (range 0-358)
Number of Pages Silent	115 (39) (range 65-156)	62.5 (43.3) (range 0-180)

Group None: no instruction in clinician's class at all

During the first half of the year when observations were collected, this table indicates that the average student received around 775 minutes (12.9 hours) of total instruction across the 52 class sessions. Recalling that the average number of classes attended was about 40, there were 19 minutes per class. Of this 19 minutes, 5.3 minutes were spent on isolated word recognition orally, 0.6 minutes on isolated word recognition silently, 4.8 minutes on oral reading of contextual mat

min

exp

typ

rec

tre

in

tac

to

var

in

two

2½

year

the

proc

the

Sec

rates

Table

material, 2.5 minutes on silent reading of contextual material, 0.8 minutes on written comprehension summaries, and 0.3 minutes of explicitly lost time. This leaves 4.7 minutes allocated to various other types of activities like class management, group instruction, and so on.

Ths basic structure of the class was oriented toward word recognition drill and contextual practice. However, within this nominal treatment was a wide variety of actual treatments. The number of minutes in class ranged from 230 to 1053 (with attendance from the previous table ranging from 18 days to 49 days in the first half year). Similar to the problems with diagnostic categorization, the problem this variation raises is the definition of equality in treatment.

Turning from treatment, the second important construct is outcome. In this analysis, two definitions of outcome will be examined: 1)difference between diagnoses at the beginning and end of the year, and 2)differences in actual performance at the beginning and end of the year. Only the computer and rule diagnoses were generated at the end of the year so the analysis could only be based on them. However, given the problems with the computerized clinician's diagnoses described earlier, the diagnoses obtained from the decision rules will be the only ones used in the analysis here.

First consider the data across treatment groups. The prevalence rates of problems with the MORAL Critical Performances are contained in Table 24.

Table 24
Prevalence Rates Before and After Treatment Across Treatment Groups

	Rule diagnosis Before Treatment	Rule diagnosis After Treatment
Instant Word Recognition	81% (34/42)	82% (28/34)
Decoded Word Recognition	90% (38/42)	53% (18/34)
Meaning Vocabulary	67% (28/42)	85% (29/34)
Oral Reading	76% (32/42)	85% (29/34)
Silent Reading	100% (42/42)	97% (33/34)
Listening Comprehension	93% (39/42)	91% (31/34)

These results seem to indicate the students still have almost as many problems as when they began. The only large difference seems to be in the prevalence of decoded word recognition problems from pre-test to post-test.

Because the students began the school year with extreme deficits, it is entirely possible that improvement did occur that was not mirrored in diagnostic classification changes. Therefore, the actual performances on the MORAL critical factors is given below:

Table 25
Distributions of Gains Across Treatment Groups

	Pre	Post	Gain
Instant Word Recognition	75.2 (24.9)	90.2 (28.4)	14.9 (11.2)
Decoded Word Recognition	8.4 (6.7)	17.3 (11.0)	8.9 (12.0)
Meaning Vocabulary	8.8 (4.9)	5.5 (4.8)	-3.5 (3.6)
Oral Reading	3.4 (1.6)	4.0 (1.9)	0.53 (1.26)
Silent Reading	4.5 (1.5)	5.9 (1.4)	1.4 (1.5)
Listening Comprehension	3.2 (1.5)	3.9 (1.3)	0.74 (1.24)

The first three rows give the numer of words recognized instantly, through mediation, and defined correctly respectively on the measurement task. The last three rows give direct grade level performances.

These raw performance data indicates that although the diagnostic classifications did not change drastically, in fact, gains did occur. Instant word recognition gained 15 words which according to the SORT description is 3/4 of a year grade equivalence. Decoding ability showed a good gain, doubling and adding almost an extra one-half year onto the students' word recognition ability. Silent reading comprehension gained a year. Realizing that these data are aggregated across treatment groups, let us now turn to the more interesting analysis of differences in treatment and gain.

Retrospective Analysis. The next clinical analysis examines cause and effect by looking backward - a retrospective analysis. Groupings must be formed on the basis of outcomes and followed backward to differences in treatment, diagnosis, or other characteristics. For this analysis, one of the major questions of interest is the relationship between various aspects of the treatment and outcome. The MORAL model of reading emphasizes the importance of relevant engaged academic time for the improvement of the reading Critical Performances.

The simplest measure of relevant engaged academic time is attendance. The 34 students remaining in the study at the end of the year were divided into various groupings depending upon their normalized gains on the MORAL Critical Performances; a score above zero was labeled high gain, a score below zero was labeled low gain. Similarly, students were separated based on high attendance and low attendance. The results are contained in Table 26.

Table 26
Differences in Rate of Poor Attenders Based on Relative Outcomes

	Relative High Gain	Relative Low Gain
Total Gain	38% (6 of 16)	59% (10 of 17)
Instant Word Recognition	27% (4 of 15)	63% (12 of 19)
Decoded Word Recognition	30% (6 of 20)	71% (10 of 14)
Meaning Vocabulary	33% (5 of 15)	61% (11 of 18)
Oral Reading	35% (6 of 17)	59% (10 of 17)
Silent Reading	44% (8 of 18)	50% (8 of 16)
Listening Comprehension	30% (7 of 23)	75% (9 of 12)

This table indicates unambiguously that gain for virtually all of the Critical Performances is dependent upon attendance. Only silent reading shows no tendency to relate attendance and gain when displayed in this way. This is most likely because very little instruction was aimed specifically at comprehension of silent contextual reading.

Given the numeric nature of both the performance gains (as opposed to diagnostic categorization gain) and the attendance, the obvious next step is to correlate attendance with performance gain. Such correlation are reported in Table 27.

Table 27
Association of Gain and Class Attendance

Category	Correlation with Attendance
Total gain	0.49
Instant Word Recognition Gain	0.29
Decoded Word Recognition Gain	0.51
Meaning Vocabulary Gain	0.03
Oral Reading Gain	0.30
Silent Reading Comprehension Gain	0.10
Listening Comprehension	0.35

Gain on decoded word recognition relative to the group demonstrates the highest correlation with attendance during the year, 0.51. In addition, the total gain across all Critical Performances is highly related to attendance. These two correlations are statistically significant at the 0.01 level. The only other correlation that almost reaches significance is listening comprehension (it is just short of being statistically different from 0 at the 0.05 level). Although the instant word recognition and oral reading correlations were <u>not</u> statistically significant, they were both a major part of instruction and are at least correlated in the right direction. Meaning vocabulary was unrelated to attendance; meaning vocabulary was not directly taught in class.

With the groups defined by their gain in the MORAL Critical Performances, correlations were computed on these groupings to see if there was any relationship between relative gain in one Critical Performance and relative gain in another. There was only one statistically significant correlation between outcome groupings (at the 0.05 level). The correlation between relative position in decoded word recognition grouping and oral reading grouping was 0.39. All other intercorrelations of grouping were 0.20 or less.

Similarly, correlations were computed on raw gain scores to see if any associations appeared there. There was only one statistically significant grouping between raw gain scores (at the 0.01 level). The correlation between decoded word recognition gain and listening comprehension gain was 0.44. All other intercorrelations were 0.30 or less.

These two correlations both involve one of the major topics of instruction in the class - decoded word recognition.

Much more data on instruction are available. Recall that during the first half of the year, observers recorded a great deal of information concerning the actual instruction that each student received besides just the attendance. The relationship between gain scores after one-half year of instruction and the specific times spent in various activities is given in Table 28.

Table 28
Associations of Outcome and Time on Class Activities

	Class time	Isolated Practice Orally	Contextual Practice Orally	Contextual Practice Silently
Total Outcome				
Good	755 (248)	223 (103)	187 (88)	91 (48)
Bad	357 (389)	90 (97)	96 (122)	51 (62)
Instant Word Recognition				
Good	795 (162)	208 (68)	181 (74)	103 (51)
Bad	449 (402)	136 (134)	127 (127)	56 (54)
Decoded Word Recognition				
Good	610 (347)	166 (112)	154 (116)	85 (58)
Bad	544 (399)	159 (127)	140 (113)	63 (56)
Meaning Vocabulary				
Good	631 (298)	190 (107)	160 (94)	76 (46)
Bad	486 (452)	123 (127)	124 (135)	68 (71)
Oral Reading				
Good	679 (341)	192 (115)	165 (112)	68 (62)
Bad	434 (376)	122 (115)	121 (111)	53 (46)
Silent Reading				
Good	654 (383)	185 (133)	177 (112)	84 (61)
Bad	511 (363)	144 (107)	122 (110)	64 (53)
Listening Comprehension				
Good	732 (298)	217 (115)	198 (95)	82 (44)
Bad	451 (386)	119 (105)	106 (111)	65 (66)

The two patterns of note are the consistent assocation of gain with time on task and the very high standard deviations on the groupings with "Bad" gain. These are both due to the retrospective technique of grouping based on outcomes, no matter what the treatment, and looking backwards to differences recorded earlier. In this case, most of the students who received no instruction, and hence zero time on all tasks, were included in the "Bad" gain groups.

Prospective Analysis. The final clinical analysis examines cause and effect by looking forward - a prospective analysis. This means that groupings must be formed on the basis of treatment and then followed to differences in outcome (T -> 0). This is by far the most straightforward analysis relating outcome and treatment. As in the previous analysis, two presentations will be given, first the differences in diagnostic categorizations across treatment groups and, then, the differences in actual raw performances.

Table 29
Prevalence Rates Across Treatment Groups

	Pretest Rule Diagnosis	Posttest Rule Diagnosis
Instant Word Recognition		
Group None	6/6	5/6
Group Half	0/4	0/4
Group Full	24/24	23/24
Decoded Word Recognition		
Group None	5/6	3/6
Group Half	3/4	2/4
Group Full	22/24	13/24
Meaning Vocabulary		
Group None	2/6	5/6
Group Half	3/4	4/4
Group Full	14/24	20/24
Oral Reading		
Group None	5/6	6/6
Group Half	0/4	0/4
Group Full	23/24	23/24
Silent Reading		
Group None	6/6	5/6
Group Half	4/4	4/4
Group Full	24/24	24/24
Listening Comprehension		
Group None	5/6	4/6
Group Half	2/4	3/4
Group Full	24/24	24/24

Looking over these data, several things are apparent. First, no student, no matter what the treatment, had much of a change in diagnostic categorization. The only significant drop in problems seems to be in decoded word recognition. In the group instructed for the full year, almost half of the deficits in decoding were eliminated. But notice that that same statement can be said about the Group None students who received no instruction!

The second point is that the clinician was clearly correct in terminating instruction for the four students in the second group. This group was different from the other students treated for the full year, even in the initial prevalence of problems. In fact, it seems questionable that these students should have been retained for instruction at all, given the data on the non treated students.

Although the instruction seemed to do little to change diagnostic categorizations, there may have been substantial change in raw performances. The differences in raw performances across treatment groups is described in Table 30.

Table 30
Performance Gains by Treatment Group

	Pretest Performance	Posttest Performance
Instant Word Recognition		
Group None	70.3 (26.1)	77.2 (29.7)
Group Half	122.8 (3.5)	142.8 (13.0)
Group Full	68.5 (17.2)	84.7 (19.6)
Decoded Word Recognition		
Group None	11.0 (6.0)	12.7 (11.5)
Group Half	13 (10.0)	18.0 (9.0)
Group Full	7.0 (6.0)	18.3 (11.3)
Meaning Vocabulary		
Group None	9.1 (5.5)	9.0 (7.8)
Group Half	11.3 (6.0)	15.0 (4.5)
Group Full	6.3 (4.0)	8.3 (4.5)
Oral Reading		
Group None	3.3 (1.8)	2.8 (1.3)
Group Half	6.3 (0.5)	7.5 (1.3)
Group Full	3.0 (1.2)	3.7 (1.5)
Silent Reading		
Group None	3.3 (1.4)	4.5 (1.4)
Group Half	3.3 (1.3)	3.8 (1.0)
Group Full	2.8 (1.3)	3.2 (1.4)
Listening Comprehension		
Group None	4.2 (1.2)	4.5 (1.4)
Group Half	4.5 (1.7)	4.8 (1.3)
Group Full	2.7 (1.3)	3.6 (1.2)

Again, the raw performance data on instant word recognition, decoded word recognition, and meaning vocabulary are numbers of words recognized or defined on the performance test. The oral reading, silent reading, and listening data are grade equivalents.

The raw performance data definitely present a different picture. First, the group of four students who received instruction for only half a year apparently learned quite a bit from the work. Their gain in instant word recognition was a full year (using the SORT conversion factor of 20 words = 1 year). Their ability to decode started out higher than any other group and still increased by 1/4 years. The oral reading ability demonstrated by this group started out three grade levels higher than the other groups - yet they showed still the largest gain - over one year. Note also that their language ability started and ended higher than any other group - almost 2 year above the initial language ability of the other instructed students. It is mainly their silent reading comprehension that is problematic.

The group of students who were instructed for the entire year also showed strong gains in word recognition. Although starting out significantly lower than the second treatment group, their instant word recognition ability was increased by 3/4 years. Also their decoding ability was raised to the same high level as the second treatment group. They were able to increase their word recognition by almost 1 year through decoding.

Finally, the no treatment group showed the lowest gains in all areas. Note that in only one Critical Performance did their abilities maintain pace with their grade placement. If the students here that received no instruction are, in fact, promoted to the next grade then they will begin that grade relatively further behind than they were the year before. The mean performance actually decreased for oral reading, probably because this performance continues to depend upon the ever more problematic word recognition ability.

Because of the small sample size, the differences in gain scores are not statistically significant at the 0.05 level with the single exception of the difference between the no instruction group and the full year of instruction group on gain in oral reading. However, the differences in initial performances of the groups are significant in a number of places.

Summary. The clinical analysis just completed raised the issue of definitions for treatment and outcome. The students in the study were divided into three treatment groups. Supposedly, the groups were to differ primarily on the amount of instruction received. However, the confounding variable of student attendance has in many cases confused this original categorization. There are students in Group Full (who supposedly received instruction all year) who in fact received less instruction than students in Group Half (who supposedly received instruction for only half a year). Clearly, actual treatment may differ substantially from intended treatment. But even ignoring the differences across treatments, the question remains as to whether or not the wide varieties of actual instruction should be treated as one treatment or as a set of closely related ones.

Similarly, the definition of outcome is in question. The traditional educational approach would be to talk of gain scores on raw performance data. Note, however, that such an analysis masks the fact that diagnostic categorization changed little.

Other Analyses

Grouping Decisions. The data from this study have so far been analyzed in the traditional methods of clinical and epidemiological research. Let us turn now to a content-area specific analysis, that of assigning students to reading groups.

The first decision made by the clinician for which the collected data can provide guidance was the selection of the 30 students to receive instruction from the tested 42. Recall that of these 42 students, four students were eliminated from consideration for non-academic reasons. This means that eight students remained to be chosen for Group None, the no instruction group. The clinician's stated policy was to choose the eight students least in need of instruction for this group. If this were the actual decision criterion used, then the normalized scores for the eight students placed into Group None should have ranked 1-8, i.e., the best eight students would not receive instruction. Table 31 indicates the degree to which this was not true.

Table 31
Decision: Choosing Students to Teach

Ideal ranking	1	2	3	4	5	6	7	8
Ranking based on Total Performance	2	4	5	7	10	11	14	24
Ranking based on Word Recognition and Silent Reading	2	3	5	8	9	17	18	20

The first row shows ideal ranking given the stated policy; the second row shows actual policy based on total diagnosis; the third row shows actual policy based on word recognition and silent reading performance. Evidently, the first half of the children actually placed into Group None were chosen properly but the second half were not.

The next decision was the splitting of the 30 students into 5 groups. The clinician, during interviews, stated that her grouping was intended to place people with the same reading level together. Rankings of total diagnoses were created for these 30 instructed students. Because of the managerial aspects of grouping decisions (e.g., conflicting class schedules), it is difficult to know how much the clinician's ideal groupings were bent by non-reading constraints. Regardless, there were six of the 30 students who were in groups substantially different from themselves (students who were the lowest or highest in their group and more than one-half standard deviations away from the next closest group member).

Both of these glimpses at the clinician's ability to informally rank student performance seem to indicate that immediate feedback of relative class standings could change some very important instructional decisions.

The only remaining concerns are those of the equivalence of alternate forms on the test battery and the reliability of various clinical decisions and actions. First, consider the equivalence of alternate forms.

Equivalence of Alternate Forms.

When the study was designed, it was decided that alternate forms of reading subtests were necessary.* Three forms of word lists and nonsense words were created by randomly reordering the words. Three alternate forms of the oral reading test (the Gray) were provided by the publisher. The design was set up so each student would use each form once. Each test session (pre-, mid-, or post-) had approximately 1/3 of the students on each form. The students were randomly assigned to the form for pretest. Midtest and posttest forms for a student were determined by progression depending upon what test a student received at pretest (1-2-3 or 2-3-1 or 3-1-2). The allocation of forms for these measures is described in Table 32.

Table 32
Distributions of Alternate Forms

Test	Pre	Mid	Post	Total
SORT1	14	14	11	39
SORT2	13	13	12	38
SORT3	15	12	11	38
GMK 1	14	14	11	39
GMK 2	13	13	12	38
GMK 3	15	12	11	38
GrayA	14	14	11	39
GrayB	13	13	12	38
GrayD	15	12	11	38

Statistical t-tests were run comparing the equivalence of these forms to one another. For the various measures pulled from the SORT

---

* For the word lists, nonsense words, and oral reading paragraphs, this now seems unnecessary given the long time span between the testing sessions and the type of performances measured. The test examiners all agreed that student recall of the test materials between testing sessions for these problem readers seemed very unlikely.

(instant word recognition, mediated recognition, total recognition, and vocabulary correctness) there were no statistically significant differences at the 0.10 level for comparisons of form 1 to 2, 2 to 3, or 1 to 3. Similarly for the Gates-McKillop there were no statistically significant differences at the 0.10 level for comparing any pair of test forms on the measures of consonant cluster recognition, phonogram recognition, nonsense word recognition, or total recognition. The Gray versions were generally similar for comparisons between any pair of test forms on the last paragraph read, average reading time, average number of miscalls, pauses, punctuation errors, repetitions, or corrections. However, two seemingly significant differences were found. When comparing form 1 to form 3 for the Gray, the average number of miscalls per selection was statistically different at the 0.10 level (mean of 8.9 compared to 7.6). When comparing form 1 to form 2, average number of punctuation errors was statistically different at the 0.01 level (mean of 0.36 compared to 0.68). Given the small actual differences and the large number of t tests run, these two differences were ignored. It was assumed, therefore, that alternate forms of these three subtests (SORT, Gates-McKillop, and Gray) were in fact equivalent.

The design also required alternate forms for the comprehension measures. Here the assumed equivalence was more problematic. The MORAL defined silent reading comprehension and listening comprehension as two different modes of input for tapping the same semantic nets in memory. Therefore, not only would three alternate forms be needed for pre-, mid-, and post-test, but another three were desired to measure silent and listening comprehension with the same measures. Remember that no

existing silent reading test with graded paragraphs and enough alternate forms could be located. Therefore, the Spache with its two forms and the Durrell with its two forms were chosen and treated as equivalent forms for measuring silent reading comprehension and listening comprehension. As it turned out, the various forms of the silent and listening comprehension measures were substantially non-equivalent. This was primarily because the two forms of the Spache were substantially different from the two forms of the Durrell.

First, the major performances of interest on the silent reading measures were the highest selection read, the average reading time, the average percentage of total possible memories during sequential recall, the average percentage correct on main idea questions, and the average percentage correct on inference questions.* These performances were substantially different when comparing the Durrell forms to the Spache. They were not all that similar when comparing the two forms of the Spache to each other or the two forms of the Durrell to each other.

Consider the measure "highest selection read." The means for these grade equivalent scores were 4.59 (s.d. 1.22) for Spache A, 4.21 (s.d. 1.52) for Spache form B, 6.13 (s.d. 1.01) for Durrell form A, and 6.17 (s.d. 1.39) for Durrell form B. The Durrell paragraphs have a mean that is statistically and meaningfully very different from that of the Spache, a difference of almost 2 years. To indicate the severity of this problem, consider the fact that all four students who were "graduated"

----

* Recall that each comprehensions test consisted of a series of selections. For each selection a student had a reading time, a percentage of possible memories, and a percentage of main idea and inference question correctness. Clearly, these values can be averaged over selections.

from the reading program at midtest time were administered a form of the Spache as a pretest comprehension measure. They scored more than two years below grade placement and were accepted into the program.

Other differences among forms are also apparent in Table 33. (Standard deviations are given in parentheses.)

Table 33
Differences in Equivalent Forms: Silent Reading

	Spache form A	Spache form B	Durrell form A	Durrell form B
highest selection read	4.59 (1.22)	4.21 (1.52)	6.13 (1.01)	6.17 (1.39)
average reading time per selection	31.96 (12.89)	34.57 (25.15)	30.58 (8.26)	31.85 (12.75)
average percentage of sequential memories	34.62 (15.81)	32.31 (12.85)	53.22 (11.92)	45.17 (11.59)
average percentage of main idea	49.27 (14.63)	39.86 (18.47)	48.25 (12.29)	44.81 (17.72)
average percentage of inference	51.79 (29.77)	59.23 (39.83)	68.52 (19.93)	56.33 (25.73)

Because of the extreme differences, the raw scores cannot be meaningfully compared. In fact it is not known at this time, what form of what test (if any) is the most accurate indication of true silent reading performance. Therefore, the following transformation will be applied to all raw data on silent reading performance.

1. Within each test session (pre-, mid-, post-) within each test form (Spache A, B, Durrell A, B), the raw scores will be standardized.

2. The standardized scores will be multiplied by the average standard

deviation of the scores across the four test forms within a test session, and added to the mean of the scores across all test forms within a test session.

This transformation assumes that the best estimate of actual silent reading performance for a particular testing session (pre-, mid-, or post-test), is the mean across all forms of a test during that test session. It further assumes that the best estimate for the standard deviation of actual silent reading performance for a particular test session is the mean of the standard deviations for the individual forms (NOT the standard deviation across all scores).

All reports of silent comprehension measures were based on the data transformed in this way.

Listening comprehension, then, had a similar problem. The performances of interest were the same as those for silent reading except that reading time was not a major concern since selections were read by the examiner not the student. Again, the performances were substantially different across the various forms of the listening comprehension data.

Table 34
Differences in Equivalent Forms: Listening

	Spache form A	Spache form B	Durrell form A	Durrell form B
average percentage of sequential memories	32.78 (11.48)	23.81 (08.42)	48.28 (11.37)	44.83 (10.50)
average percentage of main ideas	45.33 (11.53)	34.68 (16.96)	43.65 (12.17)	43.18 (11.29)
average percentage of inference	48.25 (20.96)	54.60 (17.95)	64.48 (15.88)	56.24 (15.36)

Again, the raw scores are difficult to compare meaningfully. The analogous transformation that was described for silent reading performances was applied to all raw data on listening comprehension performance. All reports of these data were based on the transformed data.

After these adjustments to the data were performed, the analysis of data based on the clinical and epidemiological research paradigm could proceed.

<u>Reliability</u> <u>of</u> <u>Data</u>.

The reliability of personal and environmental measurements, diagnostic categorizations, treatment selection and administration, and outcome measures is critical to the clinical and epidemiological research paradigms. Therefore, to as great an extent as possible, these concerns were examined here.

The first reliability analysis consisted of reliability statements on the case and environmental measures. These measures were all part of the large multi-part test battery described elsewhere. Many of the subtests had multiple items and were, therefore, amenable to reliability estimates based on homogeneity. Aside from these measures, the only other data that commented directly on the question of reliability was a test-retest reliability using pre- and post-test data. Clearly this is a less than ideal estimate because specific instruction on the tested performance occurred between the two testing sessions, but this measure of reliability might serve as a lower bound on actual reliability.

The data for these two measures of reliability on the MORAL Critical Performances raw data are described in Table 35.

Table 35
Reliability of Critical Performance Measures

	Homogeneity Estimates: Coefficient Alpha	Lower Bound on Reliability: Test-Retest Estimates: Pre-test to Post-test
Instant Word Recognition	0.95	0.92
Decoded Word Recognition	0.90	0.16
Meaning Vocabulary	-	0.73
Oral Reading	-	0.76
Silent Reading Comprehension	0.69	0.62
Listening Comprehension	0.77	0.61

These data indicate that the reliability of the word recognition measures is very high (0.9) while the reliability of connected text measures are only slightly lower (0.7). The only apparent discrepancy is the test-retest measure of decoding. In actuality, most of the students examined were brought up to some level of mastery on the decoding skills. Therefore, the alpha coefficient is a better estimate.

These data point out something else. Because the test-retest reliability is high for all other performances, the instruction during the year did little to change the students' relative position.

The second reliability concern is for reliability of diagnostic categorization. The approaches used in this study guarantee high reliability here. The decision rule diagnostic reliability and the computerized clinician's diagnostic reliability were essentially perfect; they were computerized procedures applied to raw data. The only source of error would be undetected typographical errors during data entry. The reliability of the written form diagnoses by the human clinician cannot be estimated here, but has shown to be as high as 0.7 in previous studies (Vinsonhaler, et al, in progress). For diagnostic categorization, then, the major limit on reliability seems to be the reliability of the test battery itself.

The reliability of treatment measurement was the most problematic to determine. Measures of days attended are probably quite reliable. However, the reliability of measures of time on task are questionable. In the cramped quarters of the clinician's classroom, an attempt was made to estimate inter-observer reliability of the class activity measures. Subjective impressions are that the reliability was fair, but no numbers are available. The very presence of the second observer in the room interfered with the perception of the first observer. As one observer shifted viewing from one part of the room to another, the other observer could not help but notice this and respond, thus, confusing the natural flow of the class. The only data reported here is for total time across 52 class sessions. Hopefully the variations among observers washed out with the large number of observation sessions.

The final question of reliability deals with reliability of outcome. No direct measure of this reliability was possible. However, outcome was defined as the difference between pre- and post-test diagnoses or pre- and post-test raw performances. Both of these outcome definitions rely upon other measures with good reliability. Therefore, outcome measures themselves should be reliable. As a results of this analysis, it would seem that the major parts of the diagnostic and remedial activities were reliable and the clinical and epidemiological research paradigms could be applied.

CHAPTER 5

Conclusions and Implications

The basic problem that this thesis addressed was the methodology by which clinical problem solvers try to learn from experience. The typical method for such learning is the informal aggregation of experience in which everyone naturally engages. However, such informal learning can be problematic because of biases inherent in human thinking and the structure of typical clinical experience. As a result, the quality of clinical decisions can often be substantially less than that possible through systematic examination of actual experience. In some instances, informal aggregation of experience can lead to beliefs that are in direct opposition to those actually taught by the experience.

The problems with informal learning from experience could exist in any clinical problem solving field. The field of medicine is one field in which such problems have existed in the past and still exist today. Medicine has, however, developed research paradigms that guide experience aggregation away from informal methods and more toward objective, scientific means. These paradigms are those of clinical and epidemiological research. The epidemiological paradigm is concerned with relating case (P), and environmental (E) characteristics to a diagnosis (D) while the clinical paradigm focuses on the relationships between diagnosis (D), treatment (T), and outcome (O). Both paradigms direct the careful systematic collection of data across different cases and clinicians and the analysis of the data in very specific ways. The key to these paradigms is the existence of reliable diagnostic categorizations and treatments that are differentially effective across these categorizations. Because these paradigms are based on the clinical problem solving model, it seemed likely that they could easily be generalized from medicine to other clinical fields.

As practitioners and researchers in the field of medicine continue to improve the quality of medical care, another clinical problem solving field, the diagnosis and treatment of reading problems, is just beginning to grapple with conerns of how to learn most effectvely from experience. The obvious question that was raised was whether or not the clinical and epidemiological research paradigm could be generalized to the field of reading diagnosis and, if so, what would such a generalization reveal. The purpose of this thesis, then, was first to determine how to generalize the two medical research paradigms to the field of reading and second to perform a research study guided by these paradigms searching for further understanding of reading diagnosis and treatment and for further understanding of what is required to make these formal methods of learning from experience work. Of course, it was not expected that a single such effort could rule on the adequacy of the research paradigms use in reading. However, it was expected to provide lessons in how to further examine their use in reading.

The specific clinical and epidemiological study was carried out in the classroom of one seventh grade remedial reading teacher as she examined 42 children and taught 30 of these children throughout one school year. This setting did place many limitations on the generalizability of the results. The single greatest limitation was lack of variability. All of the children observed had reading problems of some form or another; most had severe problems. The clinician assumed that a single diagnosis was adequate to describe all students, that diagnosis being that the cause of most of the reading problems was a previous instructional deficit. This led to a single treatment plan for all students, specifically, carefully sequenced practice on isolated word recognition and on contextual reading using materials with increasingly more phonetic regularities. Treatment only differed with respect to entry level and pace. These various limitations on all aspects of the clinical setting should be kept in mind as the results are reviewed.

The major clinical and epidemiological data analyses described in the previous chapter are summarized in Table 36.

Table 36
Summary of Data Analysis

	Epidemiological Analysis P,E -> D	Clinical Analysis D -> T -> O
Descriptive and Cross Sectional	*Prevalence rates of problems with Critical Performances *Agreement of different diagnostic sources across Critical Performances *Prevalence rates of diagnostic profiles of Critical Performances *Agreement of different diagnostic sources across diagnostic profiles	*Days of instruction by treatment group *Relevant engaged academic time by treatment group *Changes in prevalence rates of Critical Performance problems after treatment across treatments *Distribution of gains in Critical Performances after treatment across treatment
Retrospective	*Differences in student self perception based on diagnostic categorization *Differences in student self perception based on diagnostic categorization stratified by sex	*Differences in attendance rates based on relative outcomes *Association of gain and attendance *Association of outcome and relevant engaged academic time
Prospective	*Differences in prevalence rates of diagnostic categorizations based on ratings of language ability *Differences in prevalence rates of diagnostic categorizations based on ratings of mean length of utterance	*Differences in prevalence rate changes of Critical Performance problems after treatment by treatment *Differences in performance gains of Critical Performances after treatment by treatment

First consider the epidemiological data analysis. This analysis began with the determination of prevalence rates and profiles of problems with the Critical Performances of the MORAL. The most obvious characteristic of the data initially was that virtually all of the students examined had severe problems with all six Critical Performances, i.e., problems with instant word recognition, decoded word recognition, meaning vocabulary, oral reading, silent reading, and listening comprehension. Three different diagnostic sources were used and all three were virtually unanimous in this judgement.*

Although there was substantial agreement on this diagnostic assessment of the students, closer examination of the three types of diagnoses brought several facts to light. The three diagnostic sources generally agreed; the clinician and decision rule diagnoses were especially close. Across these three sources, however, there was an occasional sharp disagreement on the prevalence of a particular diagnostic problem. Closer inspection indicated that substantial differences in the definition of diagnostic problems existed. An example of such differences was the judgement on silent reading comprehension. The clinician and the decision rule considered a student's silent reading comprehension to be problematic if it was more than one year below grade placement; the computer simulated clinician considered it problematic if it was more than one year below word recognition level. In a very real sense, neither of these is more correct than the other. Either one of the definitions could be agreed upon as the "correct" one. What these distinctions highlight is that the precise definition of a behavior as well as a criteria for determining its adequacy are a

----

* The three types of diagnoses were (1) those generated by the reading clinician guided by a written decision aid, (2) those resulting from the computer application of simple decision rules to raw performance data, and (3) those given by a complex computer simulated clinician.

necessary prerequisite for clinical and epidemiological research. Maybe both of these diagnostic perceptions are valuable and should exist as separate diagnostic categories, but such decisions must be made consciously.

Another point of interest confirmed by the epidemiological data is the human tendency (or requirement) to simplify a problem. There were 25 different diagnostic profiles generated by the three diagnostic sources. Of these 25, the clinician used six, the decision rule used 11, and the computer simulated clinician used 13. Apparently the clinician viewed the placement of the students into diagnostic profile equivalence classes from a simpler perspective than the other two diagnostic sources. Given the 11 categories used by the rule diagnoses and the elementary and non-inferential nature of the rules, these 11 categories are probably the best estimate of the number of actual distinctions. The human clinician used half of that. Of course, one possibility is that somehow the human clinician had already performed another round of grouping and had collapsed different diagnostic profiles into treatment equivalent profiles. However, with the separate applications of a written form to each case, and with the form requiring independent judgements on each Critical Performance, this seems unlikely.

Given the reduction of the problem space for the clinician, one last implication is that actually different diagnostic classes could be incorrectly grouped together. Recalling the description in Chapter Two detailing the results of unreliable diagnoses, it is clear that such a reduction of the problem space when it is not warranted can result in errors in estimating the effect of treatment on these diagnostic categories that are nominally the same. As the clinician makes this simplification, it is possible that valuable distinctions become lost. Only more empirical study could determine if this were true.

As the study continued with the retrospective analysis of differences in personal or environmental characteristics given differences in diagnoses, several things were evident. First, for any epidemiological analysis, much more information than was collected here is required on non-problematic cases. Although in this study there were, in fact, students without a problem for any particular Critical Performance, no student examined was without problem. In this study, those that were healthy in some Critical Performance not only had problems with other ones but were also few in number. No one examined had good silent reading comprehension.

The data indicated that there was a relationship between self perception and performance. There were more students who had a correct self perception of a given Critical Performance than there were students who had an incorrect self perception. When the technique suggested from epidemiology called standardization was applied to these data using the student's sex, the resulting stratified data indicated that virtually all of the incorrect self perceptions were voiced by the male students. Whether their actual perceptions were so wrong or they felt it more necessary to put up a front is unknown. In either case, standardization proved to be a straightforward technique and apparently worked well.

The last epidemiological analysis was a prospective analysis of language ability and diagnostic categorization. Language ability was measured in two ways: after detailed observation of reading performances the examiner was asked to rate the student's language ability as high, average, or low, and to estimate the student's mean length of utterance as multiple sentence, single sentence, phrase, or word.

The association between language ability and problems with Critical Performances was pronounced. In fact, 100% of all students rated with low language ability were problematic in almost every way and further diagnoses were actually redundant. Just as important as the strong relationship between language experience and diagnosis is the comparative lack of relationship between mean length of utterance and diagnosis. The expectation had been that these two measures would give identical results. That they did not indicates that the first relationship was more than just some artifact of the rating procedure.

The clinical analysis of data began with descriptive measures of treatment and outcome across groups. The first problem raised by this analysis is similar to the problem raised by descriptive analysis of epidemiological data: problems of definition.

The students in the study were nominally receiving the same treatment within each group. In fact, however, the treatments were substantially different because of student attendance. A factor of five separated the amount of instructional time received by the lowest class attender and the highest (203 minutes to 1053 minutes during the first half term for the group instructed for the full year). Treatments with differences of this magnitude cannot be called identical treatments. Much like medical treatments which require adherance to a schedule of activities, instruction in this study required attendance. Obviously, many students could not meet this requirement. The question remains, then, when are treatments equivalent.

The second question of definition that the descriptive analysis raised concerns the definition of outcome or gain. For this sample of students across the treatment groups, very little change in diagnostic categorization occurred between pre-test and post-test. However, noticable gain did exist when the raw performance data were examined. Although not significant enough to change drastically diagnostic categorization, the passing of a year did alter actual performance.

Retrospective analysis of the clinical data indicated a strong relationship between relative class attendance and relative outcome (relative meaning compared to the rest of the 42 students). Without exception, the percentage of low attenders was smaller for high gain groupings and larger for low gain groupings. When actual correlations between raw performance gain scores and class attendance were computed, all correlations were in the right direction (more attendance give more gain); correlations for total gain and decoding skill gain with attendance were statistically different from zero. (This raises the question of whether or not patterns of data can be found to predict class attendance and used to choose which students receive some of the limited resources of remedial reading assistance.) Similar patterns were discovered when attendance was broken apart into the more fine-grained measures of relevant engaged academic time. This analysis also indicated that total time was a major determiner of total gain. Although none of the results were statistically significant, all relationships between time and gain were in the proper direction. One note: there was little relationship between silent reading gain and time on tasks, but few tasks in class were relevant to silent reading comprehension.

The final clinical analysis compared outcomes across treatment groups. This analysis indicated that treatment versus no treatment was a significant difference. The students who received no instruction averaged less than one year growth in all Critical Performances except silent reading comprehension. This means that at the beginning of next year, these students will start the year relatively farther behind than they were this year.

The differences between the groups instructed for half year and full year were also large, but not so much in terms of gain as in terms of initial diagnosis. Remember that these two groups started out together and the students instructed for the half year were "cured" in the clinician's judgement at that time. It is not surprising, then, to discover that the initial diagnoses for these students were substantially higher than for the other treated students. Perhaps the largest difference is in entering listening comprehension ability: an average of 4.5 grade equivalence for the group cured by half year and an average of 2.7 for the others.

Continuing this comparison between the two instructed groups brings out another significant fact: the final performance of the two treated groups in terms of the amount of word recognition that decoding added to their instant recognition was essentially identical. After instruction both groups were able to raise their word recognition level by almost one year with decoding. Both groups started out significantly worse than this. Such consistent gains were not seen elsewhere, so apparently, the treatment provided by the clinician is differentially effective for decoded word recognition.

Across all of the treatment groups, one last point was re-emphasized by the prospective analysis, namely, differences in gain on raw performance measures are not necessarily reflected in differences in diagnostic categorizations. Specifically, although most of the treated students did show gain, even substantial gain, ALL STUDENTS STILL HAD READING PROBLEMS. At the end of the year, every one of these students still had problems with at least one Critical Performance in reading.

The last specific analysis of data concerned the equivalence of alternate forms and the reliability of clinical measurements, decisions, and actions. Suffice it to say that problems with the actual non-equivalence of some alternate forms were accounted for and that the reliability of the personal, diagnostic, treatment, and outcome characteristics was adequate.

The final question that remains is the implications of this research for reading and its clinicians and policy makers. First of all, it appears that the question of whether or not the clinical and epidemiological research paradigms work in reading has been replaced by a number of other questions that must be answered first dealing with things like the ability to have reliable diagnostic categorizations and differential treatments. If these can be solved, it appears that the single most difficult problem in generalizing these approaches to reading may deal with the practical matters of performing such studies. It requires work on the part of clinicians and reliability at all levels. Existing microcomputer technology could assist with the careful collection of data, the statement of diagnoses, the suggestion of treatments, and the measurement of outcomes. But what benefit is provided to the clinician?

Indications from the clinician studied here are that the concerned clinician can act to make major potential improvements in treatment of problem readers given a minimal amount of formally aggregated empirical information. Long before this report was completed, our clinician altered her teaching methods given the small piece of information that oral reading gains were very slight among her students. This empirical information was in opposition to what her informal perceptions of the situation dictated. With this information, she changed instruction.

Aside from this type of information, the amount of data available to clinicians early in the development of a clinical and epidemiological data bank is small. At this time, no information is available on gains from other treatments on similar or different types of children. At this point we have data on only a very limited set of readers, with far too few good readers in the set. The question remains how the results of the data reported here would change given a larger sample of readers.

Although the data available to individual clinicians may be limited initially, as the formal data base of clinical experience builds up, many other types of assistance would be possible. The varieties and likelihoods of diagnostic profiles would be available. The empirical association of these profiles with measurement devices could lead to more reliable and valid diagnoses. Given particular diagnostic categorizations, prognoses could be determined and appropriate treatment plans suggested. As the differential effectiveness of various treatments becomes known, sequences of treatments could be prescribed, each working on a separate reading ability. Then the outcomes of each new case would serve to confirm or deny the existing practice and objective learning can occur.

This same type of information could also prove valuable to policy makers. For example, none of the children examined here had adequate silent reading comprehension at the end of the year. They will receive no special instruction next year so how will their reading ability survive? If their problem with silent reading starts to bring the other Critical Performances down with it, was it reasonable to treat them at all?

Reading is very different from medicine. For almost any medical problem, the patient has armies of microscopic allies built into the body that can battle many problems and bring the patient to full health without assistance. There is a natural state called "healthy" in medicine. In reading, it is just the opposite. If anything, the psychological aspects of a poor reader are fighting on the other side. Even the data here indicated the relationship between poor self concept and problems. Is it reasonable, then, to start curing a problem reader when one cannot complete the task?

Given empirical data on students, then, one can answer such questions. A policy maker might well decide that, with limited resources, it makes more sense to treat the readers with only moderate problems first and "save" their reading ability because it is saveable than to expend resources on children whose reading will only deteriorate again when instructional support is removed. The central concept here is to look for a change in diagnosis rather than a change in performance.

Perhaps the last statement here should be a request that researchers and clinicians in reading (and possibly any other clinical problem solving field) at least consider the possibility of developing a clinical and epidemiological research program for the formal aggregation of clinical experience. The results of such work in medicine has been impressive; the potential for similar results in reading seem promising. Although the outcomes in terms of improved clinical problem solving cannot be estimated without actually performing clinical and epidemiological research, the techniques of learning from experience by careful systematic recording and aggregation of the characteristics of that experience have an alluring common sense about them. To quote Roberts (1977, p.4) once again, clinical and epidemiological research techniques provide "possibly, the only methodological framework within which science and rationalism can be integrated into the clinical management of individual patients."

LIST OF REFERENCES

List of References

Alderson, M. An introduction to epidemiology. London: MacMillan Press, LTD, 1976.

Anderson, L.M., Evertson, C.M., & Brophy, J.E. An experimental study of effective teaching in first-grade reading groups. The Elementary School Journal, 1979, 79, 193-223.

Barnoon, S., & Wolfe, H. Measuring the effectiveness of medical decisions. Springfield, Illinois: Charles C Thomas, 1972.

Benard, J. The medical decision. In DeDombal, F.T., & Gremy, F. (Eds.),Decision making and medical care: Can information science help?. New York: North-Holland Publishing Company, 1976.

Bender, G.A. Great moments in medicine. Detroit: Parke, Davis, & Company, 1961.

Bloom, B.S. Human characteristics and school learning. New York: McGraw-Hill Book Company, 1976.

Borg, W.R., & Gall, M.D. Educational research. New York: Longman, Inc., 1979.

Brophy, J.E., & Evertson, C.M. Learning from teaching: A developmental perspective. Boston: Allyn and Bacon, Inc., 1976.

Burdette, W.J., & Gehan, E.A. Planning and analysis of clinical studies. Springfield, Illinois: Charles C Thomas, 1970.

Calfee, R. Cognitive psychology and educational practice. In D.C. Berliner (Ed.),Review of research in education, 9. American Educational Research Association, 1981.

Calfee, R., & Drum, P.A. How the researcher can help the reading teacher with classroom assessment. In L.B. Resnick & P.A. Weaver (Eds.),Theory and practice of early reading, volume 2. Hillsdale, New Jersey: Lawrence Earlbaum Associates, Publisher, 1979.

Carroll, J.B. . Teacher's College Record, 1963, 64, 723-733.

Chall, J.S. Learning to read: The great debate. New York: McGraw-Hill Book Company, 1967.

Ciba Foundation Symposium on research and medical practice: Their interaction. New York: Elsevier/Excerpta Medica/North Holland, 1976.

Coleman, V. Paper doctors: A critical assessment of medical research. London: Temple Smith, 1977.

Collen, M.F. (Ed.). Multiphasic health testing services. New York: John Wiley & Sons, 1978.

Collen, M.F., Ruben, L., Neyman, J., Dantzig, G.B., Baer, R.M., & Siegelaub, A.B. Automated multiphasic screening and diagnosis. American Journal of Public Health and the Nation's Health, 1964, 54, 741-750.

Collen, M.F., Dales, L.G., Friedman, G.D., Flagle, C.D., Feldman, R., & Siegelaub, A.B. Multiphasic checkup evaluation study: 4.Preliminary cost benefit analysis for middle-aged men. Preventative Medicine, 1973, 2, 236-246.

Cutler, P. Problem solving in clinical medicine: From data to diagnosis. Baltimore: The Williams & Wilkins Company, 1979.

Davis, F.B. The literature of research in reading with emphasis on models: Final report. East Brunswick, New Jersey: IRIS Corporation, 1970.

DeDombal, F.T., & Gremy, F. (Eds.). Decision making and medical care: Can information science help?. New York: North-Holland Publishing Company, 1976.

DeDombal, F.T., & Horrocks, J.C. Computer aided diagnosis: Conclusions from an overall experience involving 4469 patients. In J. Anderson & J.M. Forsythe (Eds.),MEDINFO 1974. New York: American Elsevier Publishing Company, Inc., 1974.

DeDombal, F.T., Leaper, D.J., Horrocks, J.C., Staniland, J.R., & McCann, A.P. Human and computer-aided diagnosis of abdominal pain: Further report with emphasis on performance of clinicians. British Medical Journal, 1974, 1, 376-380.

DeGowin, E.L., & DeGowin, R.L. Bedside diagnostic examination, third edition. New York: MacMillan Publishing Co., Inc., 1976.

Duffy, G.G. "Fighting off the alligators: What research in real classrooms has to say about reading instruction", State of the Art Conference, November 23, 1981, State University of New York at Albany.

Duffy, G.G., & Roehler, L.A. "An analysis of the instruction in reading instructional research", National Reading Conference, December 2, 1981, Dallas, Texas.

Dunkin, M.T., & Biddle, B.J. The study of teaching. New York: Holt, Rinehart, and Winston, Inc., 1974.

Einhorn, H.J. Learning from experience and suboptimal rules in decision making. In T.S. Wallsten (Ed.),Cognitive processes in choice and decision behavior. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publisher, 1980.

Einhorn, H.J., & Hogarth, R.M. Confidence in judgement: Persistence of the illusion of validity. Psychological review, 1978, 85, 395-416.

Ekwall, E. Diagnosis and remediation of the disabled reader. Boston, Massachusetts: Allyn and Bacon, 1976.

Elstein, A.S., Shulman, L.S., & Sprafka, S.A. Medical problem solving: An analysis of clinical reasoning. Cambridge, Massachusetts: Harvard University Press, 1978.

England, J.M. Medical research: A statistical and epidemiological approach. New York: Churchill Livingstone, 1975.

Feinstein, A.R. Clinical biostatistics. St. Louis: The C.V. Mosby Company, 1977.

Fisher, C.W., Filby, N.N., Marliave, R. Cahen, L.S., Dishaw, M.M., Moore, J.E., & Berliner, D.C. Teacher behaviors, academic learning time, and student achievement: Final report of phase iii-b, beginning teacher evaluation study. San Francisco: Far West Laboratory for Educational Research and Development, 1970.

Fox, J.P., Hall, C.E., & Elveback, L.R. Epidemiology: Man and disease. New York: The Macmillan Company, 1970.

Friedman, G. D. Primer of epidemiology, second edition. New York: McGraw-Hill Book Company, 1980.

Fry, E.B. Reading instruction for classroom and clinic. New York: McGraw-Hill Book Company, 1972.

Gagne, R.M. The conditions of learning. New York: Holt, Rinehart, and Winston, 1977.

Garrison, F.H. An introduction to the history of medicine, fourth edition. Philadelphia: W.B. Saunders Company, 1929.

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.

Good, T.L., Biddle, B.J., & Brophy, J.E. Teachers make a difference. New York: Holt, Rinehart, & Winston, 1975.

Good, T.L., & Brophy, J.E. Looking in classrooms. New York: Harper, Row, Publishers, 1973.

Gram, R.R. Problem solving, systems analysis, and medicine. Springfield, Illinois: Charles C Thomas, Publisher, 1972.

Hare, V.C., Jr. Systems analysis: A diagnostic approach. New York: Harcourt, Brace, & World, Inc., 1967.

Hoffmeyer, L. The process and outcome of diagnostic problem solving among reading clinicans. Unpublished doctoral dissertation, Michigan State University, 1980.

Holland, W.W. Data handling in epidemiology. London: Oxford University Press, 1970.

Holland, W.W., & Karhausen, L. (Eds) Health care and epidemiology. London: Henry Krimpton Publisher, 1978.

Keen, P.G.W., & Morton, M.S.S. Decision support systems: An organizational perspective. Reading, Massachusetts: Addison-Wesley Publishing Company, 1978.

Kurland, L.T., & Molgaard, C.A. The patient record in epidemiology. Scientific American, 1981, 245(4), 54-63.

Ipsen, J. Use of vital statistics. In W.W. Holland & L. Karhausen (Eds.),Health care and epidemiology. London: Henry Krimpton Publishers, 1978.

Jackson, P.L. Life in classrooms. New York: Holt, Rinehart, and Winston, Inc., 1968.

Lange, H.J. Decision-making by doctors aided by computers. In F.T. DeDombal & F. Gremy (Eds.),Decision making and medical care: Can information science help?. New York: North-Holland Publishing Company, 1976.

Leaper, D.J., Horrocks, J.C., Staniland, J.R., & DeDombal, F.T. Computer-assisted diagnosis of abdominal pain using "estimates" provided by clinicians. British Medical Journal, 1972, 4, 350-354.

Leibert, R.E. (Ed.). Diagnostic viewpoints in reading. Newark, Delaware: International Reading Association, 1971.

Lilienfeld, A.M., & Lilienfeld, D.E. Foundations of epidemiology, second edition. New York: Oxford University Press, 1980.

Lilienfeld, D.E., & Lilienfeld, A.M. Epidemiology: A retrospective study. American Journal of Epidemiology, 1977, 106, 445-459.

Lusted, L.B. Introduction to medical decision making. Charles C Thomas, Springfield, Illinois, 1968.

Mausner, J.S., & Bahn, A.K. Epidemiology: An introductory text. Philadelphia: W.B. Saunders Company, 1974.

Mead Johnson Strategies in clinical research. Evansville, Indiana: Mead Johnson, 1976.

Neustadt, R.E., & Fineberg, H.V. The swine flu affaie: Decision making on a slippery disease. U.S. Department of Health, Education and Welfare, 1978.

Nossal, G.J.V. Medical science and human goals. London: Edward Arnold, 1975.

Peterson, P.L., & Walberg, H.J. Research on teaching: Concepts, findings, and implications. Berkley, California: McCutchan Publishing Corporation, 1979.

Roberts, C.J. Epidemiology for clinicians. Tunbridge Wells: Pitman Medical Publishing Company, 1977.

Simon, A., & Boyer, E. Mirrors for behavior: An anthology of observation instruments. Philadelphia: Research for Better Schools, Inc., 1967.

Simon, A., & Boyer, E. Mirrors for behavior: An anthology of observation instruments continued, 1970 suppliment, volumes a and b. Philadelphia: Research for Better Schools, Inc., 1970.

Sherman, G.B. Untitled manuscript. In preparation, 1982-1983.

Siegelaub, A.B., Friedman, G.D., Collen, M.F., & Kodlin, D. Research applications of MHTS. In M.F. Collen (Ed.),Multiphasic health testing services. New York: John Wiley & Sons, 1978.

Singer, H., & Ruddell, R.B. Theoretical models and processes of reading, second edition. Newark, Delaware: International Reading Association, 1976.

Small, W.P., & Krause, U. An introduction to clinical research. Baltimore: The Williams & Wilkins Company, 1972.

Smith, F. Understanding reading: A psycholinguistic analysis of reading and learning to read. New York: Holt, Rinehart, & Winston, Inc., 1971.

Spache, G.D. Diagnosing and correcting reading disabilities. New York: Allyn & Bacon, Inc., 1976.

Spache, G.D. Investigating the issues of reading disabilities. New York: Allyn & Bacon, Inc., 1976.

Spache, G.D., & Spache, E.B. Reading in the elementary school, third edition. New York: Allyn & Bacon, Inc., 1973.

Stallings, J.A. Learning to look: A handbook on classroom observation and teaching models. Belmont, California: Wadsworth Publishing Company, Inc., 1977.

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.

Sullivan comprehension readers. Palo Alto, California: The Learning Line, 1980.

Tversky, A., & Kahneman, D. Judgement under uncertainty: Heuristics and biases. Science, 1974, 185, 1124-1131.

VanRoekel, J. The problem solving behavior of reading specialists and learning disabilities resource teachers. Unpublished doctoral dissertation in progress, Michigan State University, .

Vinsonhaler, J.F., Weinshank, A.B., Wagner, C.C., & Polin, R. Diagnosing children with educational problems: Characteristics of reading and learning disability specialists and classroom teachers. Reading Research Quarterly, In preparation, due approximately fall, 1982.

Vinsonhaler, J.F., Weinshank, A.B., Polin, R., & Wagner, C.C. Model based training to improve diagnostic reliability in reading. Planned for Reading Research Quarterly, In preparation, due approximately spring, 1983.

Vinsonhaler, J.F., Weinshank, A.B., Cureton, D., & Blatt, G. A model of reading and learning to read. Unpublished manuscript, 1980.

Weinshank, A.B. An observational study of the relationship between diagnosis and remediation in reading. Unpublished doctoral dissertation, Michigan State University, 1980.

Wilson, P.D., Horrocks, J.C., Lyndon, P.J., Yeung, C.K., Page, R.E., & DeDombal, F.T. Simplified computer-aided diagnosis of acute abdominal pain. British Medical Journal, 1975, 2, 73-75.

Wulff, H.R. Rational diagnosis and treatment. Oxford: Blackwell Scientific Publications, 1976.

Yerushalmy, J. The statistical assessment of the variability in observer perception and description of roentgenographic pulmonary shadows. In W. Winkelstein Jr. & F.E. French (Eds.),Basic readings in epidemiology. New York: MSS Information Corporation, 1972.

References: Reading Tests

Durrell, D.B. Durrell analysis of reading difficulty: new edition. Harcourt, Brace, & World, Inc., 1955.

Gates, A.J. & McKillop, A.S. Gates-Mckillop reading diagnostic tests. Teachers College Press, 1962.

Gray, W.S. Gray oral reading tests. Bobbs-Merrill Company, Inc., 1967.

Slosson, R.L. Slosson oral reading test (SORT). Slosson Educational Publications, Inc., 1963.

APPENDICES

Appendix A: Test Battery Stimulus Materials

The following pages include the stimulus materials for the measurement device. The students in the junior high school class used in the study had these materials in front of them, covered in plastic and in a three ring notebook. The notebook contained materials for all alternate forms of each subtest. The test examiner would direct the student to appropriate sections of the notebook. The notebook contains 43 pages. The contents presented in this appendix are identical to that used in the study except that the format has been altered slightly.

1. jump
2. want
3. look
4. play
5. come
6. is
7. can
8. one
9. see
10. ball
11. little
12. three
13. baby
14. run
15. down
16. mother
17. here
18. up
19. help
20. make
21. very
22. under
23. hill
24. road
25. dark
26. horse
27. came
28. puppy
29. was
30. bump
31. with
32. ride
33. along
34. friends
35. first
36. food
37. wish
38. what
39. basket
40. live
41. better
42. hide
43. large
44. field
45. around
46. suddenly
47. forest
48. river
49. sheep
50. station
51. breakfast
52. across
53. heavy
54. lunch
55. happen
56. grass
57. game
58. farmer
59. stars
60. hope

61. hunger
62. evening
63. bench
64. excuse
65. grove
66. desire
67. perform
68. destroy
69. delicious
70. understood
71. safe
72. against
73. timid
74. ocean
75. damp
76. stream
77. empty
78. smash
79. stone
80. reward
81. claimed
82. forehead
83. serious
84. speechless
85. distant
86. dainty
87. anger
88. courage
89. slumber
90. common
91. harness
92. flakes
93. develop
94. vacant
95. price
96. future
97. appearance
98. promptly
99. region
100. silence
101. dignity
102. generally
103. gracious
104. applause
105. define
106. cushion
107. jungle
108. haze
109. profitable
110. tailor
111. fragrant
112. obedient
113. marriage
114. presence
115. terrace
116. interfere
117. ambition
118. merchant
119. custom
120. extended

121. tremendous
122. dungeon
123. customary
124. abundant
125. malicious
126. imaginary
127. excellence
128. consequently
129. installed
130. responsible
131. yearning
132. liquid
133. medicine
134. spectacular
135. rebellion
136. compliments
137. infected
138. inventory
139. importance
140. detained
141. reminiscence
142. architecture
143. environment
144. counterfeit
145. pensive
146. tremor
147. intricate
148. administer
149. continuously
150. attentively
151. industrious
152. contemporary
153. approximate
154. standardize
155. complexion
156. exhausted
157. malignant
158. compassionate
159. crisis
160. society
161. nucleus
162. proportional
163. articulate
164. remarkably
165. nonchalant
166. memorandum
167. deprecate
168. contrasting
169. whimsical
170. evident
171. prairies
172. twilight
173. intangible
174. supplement
175. formulated
176. exuberant
177. antique
178. inducement
179. irrelevance
180. grotesque
181. sojourn
182. affable
183. auspicious
184. enamoured
185. futility
186. primordial
187. inadequacy
188. chastisement
189. envisage
190. pandemonium
191. panorama
192. compressible
193. gustatory
194. traverse
195. simultaneous
196. excruciating
197. contraband
198. decipher
199. scrupulous
200. facsimile

1. one
2. little
3. ball
4. play
5. down
6. mother
7. see
8. run
9. baby
10. jump
11. look
12. make
13. three
14. is
15. here
16. help
17. come
18. want
19. can
20. up
21. along
22. with
23. bump
24. hill
25. wish
26. under
27. horse
28. came
29. very
30. food
31. puppy
32. was
33. basket
34. ride
35. road
36. live
37. first
38. what
39. dark
40. friends
41. lunch
42. hide
43. field
44. heavy
45. suddenly
46. across
47. river
48. hope
49. game
50. sheep
51. breakfast
52. happen
53. grass
54. around
55. forest
56. station
57. stars
58. better
59. large
60. farmer
61. stone
62. against
63. ocean
64. grove
65. smash
66. stream
67. damp
68. safe
69. excuse
70. understood
71. perform
72. empty
73. hunger
74. reward
75. destroy
76. desire
77. delicious
78. evening
79. bench
80. timid
81. forehead
82. serious
83. claimed
84. flakes
85. slumber
86. promptly
87. courage
88. harness
89. distant
90. vacant
91. anger
92. dainty
93. develop
94. appearance
95. silence
96. future
97. speechless
98. region
99. price
100. common
101. fragrant
102. tailor
103. interfere
104. terrace
105. generally
106. define
107. custom
108. ambition
109. dignity
110. applause
111. extended
112. merchant
113. haze
114. marriage
115. cushion
116. jungle
117. gracious
118. profitable
119. presence
120. obedient

121. malicious
122. yearning
123. inventory
124. installed
125. customary
126. consequently
127. detained
128. medicine
129. dungeon
130. compliments
131. rebellion
132. imaginary
133. excellence
134. abundant
135. spectacular
136. infected
137. liquid
138. importance
139. responsible
140. tremendous
141. standardize
142. counterfeit
143. tremor
144. complexion
145. environment
146. approximate
147. exhausted
148. attentively
149. pensive
150. contemporary
151. industrious
152. administer
153. compassionate
154. architecture
155. reminiscence
156. crisis
157. continuously
158. malignant
159. intricate
160. society
161. evident
162. proportional
163. contrasting
164. memorandum
165. grotesque
166. prairies
167. antique
168. formulated
169. articulate
170. supplement
171. twilight
172. nonchalant
173. whimsical
174. intangible
175. inducement
176. exuberant
177. deprecate
178. remarkably
179. nucleus
180. irrelevance
181. primordial
182. futility
183. enamoured
184. decipher
185. chastisement
186. traverse
187. contraband
188. compressible
189. scrupulous
190. auspicious
191. simultaneous
192. inadequacy
193. panorama
194. excruciating
195. sojourn
196. envisage
197. affable
198. facsimile
199. pandemonium
200. gustatory

1. play
2. see
3. make
4. three
5. help
6. baby
7. down
8. is
9. can
10. up
11. one
12. run
13. want
14. here
15. ball
16. jump
17. little
18. come
19. mother
20. look
21. road
22. basket
23. live
24. very
25. bump
26. ride
27. first
28. wish
29. what
30. horse
31. along
32. dark
33. came
34. with
35. under
36. food
37. hill
38. puppy
39. friends
40. was
41. station
42. breakfast
43. sheep
44. field
45. better
46. heavy
47. around
48. river
49. forest
50. suddenly
51. lunch
52. happen
53. large
54. hope
55. grass
56. farmer
57. hide
58. stars
59. game
60. across
61. stream
62. desire
63. evening
64. understood
65. destroy
66. delicious
67. ocean
68. grove
69. excuse
70. damp
71. hunger
72. bench
73. empty
74. timid
75. smash
76. perform
77. stone
78. safe
79. against
80. reward
81. dainty
82. price
83. courage
84. develop
85. forehead
86. serious
87. flakes
88. distant
89. appearance
90. anger
91. promptly
92. speechless
93. common
94. slumber
95. region
96. future
97. harness
98. claimed
99. vacant
100. silence
101. extended
102. define
103. profitable
104. obedient
105. tailor
106. applause
107. gracious
108. marriage
109. interfere
110. generally
111. presence
112. jungle
113. terrace
114. ambition
115. custom
116. cushion
117. fragrant
118. haze
119. merchant
120. dignity

121. responsible
122. malicious
123. detained
124. installed
125. inventory
126. imaginary
127. excellence
128. spectacular
129. abundant
130. tremendous
131. medicine
132. consequently
133. yearning
134. importance
135. compliments
136. dungeon
137. liquid
138. customary
139. infected
140. rebellion
141. complexion
142. crisis
143. approximate
144. tremor
145. intricate
146. compassionate
147. contemporary
148. society
149. standardize
150. exhausted
151. malignant
152. reminiscence
153. industrious
154. pensive
155. architecture
156. environment
157. continuously
158. attentively
159. administer
160. counterfeit
161. prairies
162. nucleus
163. antique
164. contrasting
165. twilight
166. articulate
167. memorandum
168. supplement
169. deprecate
170. exuberant
171. grotesque
172. irrelevance
173. nonchalant
174. proportional
175. formulated
176. whimsical
177. remarkably
178. inducement
179. evident
180. intangible
181. envisage
182. pandemonium
183. futility
184. affable
185. contraband
186. panorama
187. traverse
188. enamoured
189. compressible
190. excruciating
191. inadequacy
192. chastisement
193. auspicious
194. decipher
195. primordial
196. simultaneous
197. sojourn
198. facsimile
199. gustatory
200. scrupulous

1.sk
2.fr
3.st
4.sl
5.ch
6.cr
7.gl
8.gr
9.sm
10.dw
11.pr
12.cl
13.wr
14.th
15.tw
16.sw
17.wh
18.bl
19.sh
20.dr
21.fl
22.tr
23.pl
24.idge
25.er
26.ate
27.ite
28.ind
29.een
30.ible
31.ock
32.ome
33.ell
34.emp
35.ickle
36.ed
37.ew
38.ick
39.ark
40.asp
41.oy
42.ight
43.able
44.ing
45.ow
46.ack
47.stind
48.smow
49.chible
50.whickle
51.proy
52.dwer
53.twasp
54.shable
55.flark
56.thate
57.slome
58.trock
59.swite
60.gright
61.plew
62.glemp
63.clidge
64.bleen
65.drack
66.sked
67.wrick
68.fring
69.crell

1.st
2.dr
3.ch
4.tw
5.gr
6.wr
7.dw
8.cl
9.bl
10.th
11.wh
12.sm
13.sw
14.sl
15.sh
16.fr
17.cr
18.tr
19.pl
20.fl
21.pr
22.sk
23.gl
24.ight
25.idge
26.emp
27.er
28.ack
29.ite
30.een
31.ark
32.ickle
33.oy
34.ew
35.asp
36.ick
37.ed
38.ind
39.ome
40.ing
41.ock
42.ate
43.ible
44.ow
45.able
46.ell
47.flark
48.chible
49.smow
50.clidge
51.dwer
52.sked
53.thate
54.gright
55.bleen
56.glemp
57.stind
58.wrick
59.whickle
60.crell
61.swite
62.twasp
63.plew
64.proy
65.trock
66.drack
67.shable
68.fring
69.slome

1.sl
2.tr
3.gl
4.th
5.tw
6.wh
7.sw
8.ch
9.cr
10.pl
11.sm
12.sh
13.wr
14.sk
15.bl
16.fl
17.cl
18.st
19.dr
20.gr
21.pr
22.dw
23.fr
24.oy
25.idge
26.ick
27.ate
28.able
29.ing
30.ome
31.ow
32.er
33.asp
34.ack
35.ind
36.ight
37.ible
38.een
39.ock
40.ite
41.ell
42.ark
43.ickle
44.ew
45.ed
46.emp
47.shable
48.twasp
49.plew
50.glemp
51.drack
52.slome
53.smow
54.whickle
55.bleen
56.crell
57.gright
58.thate
59.swite
60.clidge
61.fring
62.flark
63.wrick
64.stind
65.proy
66.dwer
67.chible
68.trock
69.sked

A3

One morning a boy made a boat. "Where can I play with it?" he asked. Father said, "Come with me in the car. We will take your boat with us." Soon the boy called, "Please stop. I see water. May I play here?" "Yes," said Father. "Have a good time."

A4

One day five children went out to play in the beautiful white snow. They played for a long time and then began to make snow animals.

One of the animals was a dog. Soon the dog next door came out of the house. When he saw the snow dog he said, "Bow-wow."

The children laughed. "Now we have a dog that can bark."

A5

It was pet day at the fair. The children were waiting for the parade of animals to begin. They had trained their pets to do many different tricks. Among them was a tall boy whose goat made trouble for him. It kicked and tried hard to break away. When it heard the band, it became quiet. During the parade it danced so well that it won a prize.

A6

Airp

freight, a

dangerous

or herds a

our zoos.

highways.

A7

Hundre

China, a la

rich, civi

distant re

languages

years.

A8

The e

eight hard

the much f

The specta

faces were

A6

Airplane pilots have many important jobs. They fly passengers, freight, and mail from one city to another. Sometimes they make dangerous rescues in land and sea accidents, and drop food where people or herds are starving. They bring strange animals from dense jungles to our zoos. They also serve as traffic police and spot speeding cars on highways.

A7

Hundreds of years ago, most of Europe was a very poor region. But China, a large country in eastern Asia, had many of the comforts of a rich, civilized nation. Only a few people from Europe had visited this distant region. One was the famous Marco Polo. He learned some of the languages that were spoken in China and served its great ruler for many years.

A8

The eager spectators who had cheered the plucky Warriors through eight hard-fought innings were silent. Only a run was required to defeat the much feared champions, who had previously defeated all opponents. The spectators had earlier criticized the umpire severely. Now their faces were tense with excitement as the players took their positions.

A9

The oil industry has been greatly increased by recent advances in science. Geologists have discovered new ways of locating veins of oil-producing rock. Problems of gusher control have been solved. Very effective also are newer methods of refining crude oil which have resulted in a higher ratio of quality fuel oil from a given volume of crude oil.

A10

In response to the impulse of habit Joseph rose and spoke as in former days. He spoke vigorously, continuously, and persuasively while the others listened attentively but in grim and contemptuous silence. Finally exhausted, Joseph hesitated for a moment; as often happens in such circumstances he became confused and was unable to resume speaking.

A11

Many of the hypotheses about physical phenomena formulated by early philosophers were inconsistent and in most cases could not be universally applied. In order to develop accurate principles very capable physicists, mathematicians, and statisticians had to cooperate wholeheartedly over long periods of time to verify numerous basic facts and assumptions.

A12

In a

contraste

the eartl

distance

students

expressio

A13

Duri

clambered

a canopy

beauty nc

mortar s

calm and

A12

In a concluding lecture on sidereal spaces, the astronomer contrasted the infinitesimal difference in the distance of the moon from the earth at apogee and at perigee with the great difference in the distance of the earth from the sun at aphelion and at perihelion. The students interrogated him, evidencing precociousness and lucidity in expression.

A13

During a hiatus in the desultory firing, the apt lieutenant clambered wearily over the detritus piled against the redoubts. Beneath a canopy of empyrean blue lay the quiet, bucolic landscape, its pristine beauty now defiled by myriad diminutive promontories thrown up by the mortar shells, but radiating momentarily an inexplicable if spurious calm and peace.

B3

A bo

with a bo

home." In

Three kit

B4

A lit

she said,

Mothe

had a big

hat. It wa

B5

Twelv

picture of

paper bags

tried to p

legs, head

prize.

B3

A boy had a wagon. He ran with it to a store. Soon he came back with a box. He called his dog and said, "Jump in, Happy. This is your home." In the morning the boy had a surprise. Happy was not in the box. Three kittens were there.

B4

A little girl ran out of a white house into a big yard. "Mother," she said, "my pet bird is gone. It went out of the open window."

Mother laughed and said, "Look on my hat." When the girl looked she had a big surprise. A yellow bird with blue wings was on Mother's pretty hat. It was the bird that flew away.

B5

Twelve boys were waiting in line at a party to play a game. A picture of a lion hung on the wall before them. They first put large paper bags over their heads so they couldn't see. Each of the boys then tried to pin a ribbon on the lion's tail. They put ribbons on the lion's legs, head, and body. All missed its tail. So none of them won the prize.

B6

One morning a big poster outside of Oak School told people about a basement bargain sale. Inside were long counters on which things collected by the children were displayed. Price tags were fastened to all articles. Most of the customers bought old but useful furniture. The sale was a huge success, and the money was used to purchase library books.

B7

All of us admire the great skill of a good truck driver. He hauls many tons of things almost daily, including dangerous explosives. On mountain roads and in other isolated places he faces real dangers alone. He is his own mechanic. Sturdy and dependable, he will interrupt his schedule to help anyone who encounters real difficulty on a highway.

B8

Rocky portions of the earth's surface are always changing. Many huge glaciers in the mountains carry along immense boulders which crush the rocks beneath. Chemicals in many streams penetrate rocks and dissolve them. Rocky surfaces are also broken up by processes of freezing and thawing which occur in most regions of alternate hot and cold weather.

B9

After the American Revolution the colonies became states, each one having a governor. What was urgently needed was a federal government to insure domestic peace and to protect citizens from enemy attack. A constitutional convention was convened. After heated controversy, a constitution was prepared and submitted to the states for approval.

B10

Beside the fireplace with its polished fixtures was a mohair chair which was in sharp contrast with a brilliant cover on a near-by footstool. Against the opposite wall stood a desk with stationery protruding from all its pigeonholes. But the object to which Alice's eyes returned repeatedly was a large flagon of incomparable value and startling beauty.

B11

The visage of the pontiff was a familiar sight amidst the ornate decorations of the court. Famous for his politeness, he was as familiar with worldly affairs as with theology--a master strategist who could mold saints and sinners into a unified group or, if the situation justified such steps, discountenance unregenerates with a single quiet reproof.

B12

An immediate rejection of customary rituals was unlikely as cultists steeped in traditional tribal lore advocated propitiation of imaginary deities. Their stubborn opposition to the abandonment of paganism subsequently brought opprobrium upon them, and historians record a cessation of eleemosynary enterprises until such customs were discontinued.

B13

The ophthalmologist sent cultures to the microscopist requesting his opinion as to the causative organism of a painful conjunctivitis. A delay resulted when the expert on microscopy consulted the histomorphologist before giving a report. Meanwhile an anodyne was prescribed to forestall a recurrence of the patient's unfortunate sharp painful attacks.

D3

A cat wanted to find her kittens. She looked in the house and all over the farm. But she could not find them. Soon the mother cat saw a girl. "Mew," she said. "Help me find my kittens." "Look," laughed the girl. "Your kittens are coming to find you."

D4

One Saturday a black dog was walking down a pretty city street. He had no home and no one to feed or to pet him. At last he was so tired he lay down to sleep for a short time.

Soon a small boy came by. "Are you lost?" he said. "Do you have a good home? I am glad I found you because I want a dog for a pet."

D5

One bright summer day twin brothers walked to a lake with their uncle to fish. They sat still for a long time waiting for the fish to bite. Finally one boy got a bite. He became so excited that he dropped his pole into the water. The fish quickly swam away with it. Soon the pole disappeared. The surprised boy looked at his uncle and then laughed.

D6

All over the world farmers face many difficult problems. Insects, birds, and other animals attack most of the farm crops and harm growing vegetables and fruit. At any season of the year the weather may also damage crops. Extreme heat, cold, or rain may ruin fruit, vegetables, or grain. However, most farmers usually succeed in having excellent crops.

D7

Both city and country children love to visit the zoo. A child who likes animals often finds all his familiar outdoor friends there. The fat woodchuck waddles as he always does in his native home. The playful seals swim in pools of cool water, flapping their flippers rapidly. Polar bears walk slowly back and forth paying little attention to people.

D8

An airplane pilot eagerly hoped for a career in the Planet Air Force. His own occupation was very crowded and he was ambitious. He did not anticipate with pleasure the longer flight schedules, but the new position provided opportunitv for advancement. He might sometime even direct maneuvers of the Flying Demons, the "mounted police" of outer space.

D9

At last the captain had selected a sturdy vessel for his dangerous voyage and prepared to sail. His first mate was a reformed buccaneer who was a proud but clever knave, and a deadly opponent in hand-to-hand encounters. The crew too was a surly lot, but the best he could get on short notice. Even before they embarked he saw mutiny lurking in their eyes.

D10

At first the energetic monk was heedless to the ferocity of the storm. He even accepted cheerfully nature's challenge to his fortitude. But soon the wind's incessant rasping in the bushes lessened his enthusiasm, and by the time he was within a furlong of the secluded monastery, he was anticipating eagerly temporary respite from the fury of the gale.

D11

When we returned from the scenes of Asia Minor and the Grecian Archipelago, our ecstatic praise of the architecture of ancient Greece was equaled only by our laudations of the many archeological wonders of Egypt. We confess preference for the classical antiquities which evoke a mystic calm and imbue our inept minds with great reverence for the Acropolis.

D12

Disconcerted by the spread of Origenistic opinion in Syria, Justinian condemned fourteen propositions espoused by The Alexandrian and compelled the Church to continue condemnation of his theories and to excommunicate Origen. Later, through the machinations of certain prelates a lengthy and mischievous controversy over the Monophysitic heresy was provoked.

D13

The areas of regional metamorphism on the Meseta are formed of plutonic rocks passing upwards into a strata crystalline series in the Central Cordillera range. This series shows three levels: the lowest, augen-gneiss; an intermediate one, micaceous gneiss with schists and crystalline limestones introduced; and the upper one, lustrous schists and phyllites.

1. Sam is wet. A bat can fly. Ned sat up in bed. A fox bit Tim on the leg.

2. Bob and Pam sat on a raft. Bob had a fish in his hand. Pam had a ship.

3. Frank sent Nan a gift. She gives him a kiss to thank him.

4. A shark is a big fish that can harm a man. The man swims as fast as he can.

5. Bill has come in first. He has won. Jack is last. He has lost.

6. This bride has a long gown on. She stands in front of a church.

7. We can play lots of games with cards. These three cards are the ace of clubs, the king of clubs, and the queen of spades.

8. I eat three meals a day. For lunch I had a pound of ground meat, a slice of toast, and a cup of tea.

9. This nice young man has lost his sight. He is now blind, but he has a smart dog for a guide.

10. The player who throws the ball is the pitcher. The one who swings at it is the hitter. The man who crouches in back of the plate is the catcher.

11. Mister Smith is a father. His wife is a mother. They are both taller than their son or their daughter.

12. Jan White is washing her clothes in the washer and drying them in

the drier. When she is through, she will hang them up on the hangers.

13. The boy and the girl are seated at their desks. The boy is reading a book. The girl is writing a letter.

14. Winter is the coldest time of the year. The skies are gray. The days are cloudy, windy, and snowy.

15. In the brightly lit big tent, a monkey is riding on the back of a donkey. Each child has money in his hand. Billy has a quarter and Jimmy has a dollar bill.

16. The boy wearing the mask is trying to frighten the men and children. The puzzled poodle is barking at what he thinks is a witch.

17. This woman has a bonnet on her head and a bracelet on her left wrist. She is sitting at the window and feeding pretzels to a sparrow, a pigeon, and a squirrel.

18. In backing her car out of the garage, a woman driver hit a hydrant. Her luggage spilled out of the trunk in front of a cyclist who stopped all the traffic on the street.

19. The youthful rider is mounted on a prancing stallion. In the flat prairie below, a train is just pulling out of the station. To the far right, storm clouds are forming over the tall mountain peaks.

20. In the foreground, a barefoot woman in a bathrobe is using clothespins to pin up a bedspread. To her left, her younger brother is beating the dust out of a mattress. In the background, her uncle

is painting a bookcase.

2a

Bob has a little red wagon. He likes to ride in it. He pulls it slowly up the hill. Then he rides it quickly down again.

One day he took his dog with him. He pulled the dog up the hill. Then they rode down the hill. But the dog did not like to ride down. He jumped out of the red wagon. Bob went down by himself.

Now he does not try to take his dog in the wagon.

3a

Bob has a brown and white dog named Spotty. He is called Spotty because he has brown spots on his nose. Bob always takes his dog on his trips to the woods. The dog helps scare the rabbits. Bob walks slowly, but his dog scampers through the leaves.

One day Spotty left Bob and went off by himself. Bob called and whistled, but the dog did not come back to him. After a while Bob heard the dog barking a long way off. Bob walked toward the sound of the barking until he found the dog. Spotty thought he had caught a black and white kitten. But it wasn't a kitten, it was a skunk. That night the dog had to sleep outside.

4a

Yesterday Bob took a trip to a city market that was somewhat like a store but a great deal bigger. It didn't have any bread or canned goods like the grocery stores. But there were a great many big boxes of vegetables and fruits.

Bob was hungry and wanted just one plum or cherry to taste. He wondered if one of the men would sell him just one plum. Everyone was buying the fruit and vegetables by the whole crate. When Bob asked the man to sell him one plum, he laughed and gave Bob an extra large plum wrapped in paper but wouldn't take any money.

As he walked along eating the plum, Bob watched the men unloading the trucks and big trailers. They would chop open the top of the crate so that anyone could see the fruit. If a buyer liked the fruit, and was willing to pay the price, he might buy the entire truckload.

5a

As a ship's boy, John Paul had all sorts of odd jobs on board. Sometimes he scrubbed decks or helped the cook. He cleaned the captain's cabin and ran errands, but he had other duties that pleased him more. He helped to clean the guns, which the merchant ship carried for protection. And several times he stood behind the big wheel to steer the ship.

Captain Benson wrote in the ship's log, or daily record, that the trip was calm and smooth-sailing. Nothing unusual happened, but every day was a real adventure for the new ship's boy. At the end of the voyage it was a thrill to sight land. When the ship docked near Fredericksburg, Virginia, John Paul was waiting to go on shore.

John Paul's brother had a tailor shop in Fredericksburg and was very happy in his new home. He was eager to talk about the wonderful country, but John Paul already loved America. During the next few years John Paul visited America often. He became used to the free and democratic ways of the new country. Meanwhile he had learned to be an expert sailor. Although he was not tall, he was strong and quick. With his long arms he could haul or trim a sail with the best of men.

6a

Elephants are found wild today only in warm regions--in tropical Africa and in India. The story was very different 50 thousand years ago. Then, two species of the elephant family roamed North America and Europe in vast numbers.

One of them was the mastodon. The mastodon lived in the eastern part of our country during the period of the Great Ice Age. In the swamps that were formed when the ice disappeared, many of the huge creatures were trapped and killed. We have found some of their skeletons. At a glance, the mastodon must have looked much like the elephants of today, except that it was covered with coarse, woolly hair and its tusks were much larger. It was probably heavier than the elephants we know, but not taller. Its head was flatter and its lower jaw longer. Its teeth were not like the teeth of the elephants of today.

More than 200 years ago, the people of New England found bones of the mastodon when they dug ditches to drain swamps. At first they thought that the bones they found were bones of giant people. When they found teeth that weighed more than four pounds apiece, they decided that the giants were giants indeed.

7a

Just as in driving a car, we use at least three speeds in reading. High gear in reading is called skimming, while studying is reading in low gear. Between these two, at a second gear, is what might be called a moderate speed of reading. As you may have heard, the good reader adapts his rate to the purpose of his reading. The rate he uses is determined by how much he wants to get out of the material he is reading. His rate is also influenced by the difficulty of the reading material. Thus, he shifts from gear to gear according to the amount he wants to retain or how difficult he finds the going.

Skimming is useful for a number of situations in reading. We can use it when looking for a particular fact on a page or in a table. It is also appropriate when we have to cover a large amount of material that is not too interesting or too important. Skimming may also be used to determine the general trends or ideas of a selection when we do not have to know the fine details. It is also helpful when we are making a quick brush-up before recitation. Finally, it is very useful as the speed at which we would do pre-reading before studying intensively.

8a

President Thomas Jefferson, in 1804, commissioned an expediton to go into the Northwest Territory to explore the land that was bought in the now-famous Louisiana Purchase. Meriwether Lewis, private secretary to the President, and Captain William Clark of the United States Army headed the 26 men who started up the Missouri River from St. Louis on May 21, 1804. On July 18 of of that year the group reached the southwest corner of the present state of Iowa, proceeded northward along the Missouri, and traversed parts of Iowa many times. The Lewis and Clark State Park west of Onawa was named in honor of these explorers.

Sergeant Charles Floyd, a Kentucky backwoodsman and one of the most competent men of the party, became ill on August 19, 1804, and died the following day. His body was laid to rest upon a high bluff near the present site of Sioux City, where it is marked with a tall monument. The Floyd River and Sergeant Bluff were named in his honor.

The expedition proceeded westward to the mouth of the Columbia River, and returned to Washington during the early months of 1807. Lewis was appointed governor of the Louisiana Territory and Clark was named governor of the Northwest Territory.

2b

Bob was eating his dinner. "Ding-dong," went the fire bell. The firemen were going to a fire. The engine went quickly down the street.

Bob jumped up and left his dinner. He ran into the street.

The fire was on the next corner. Bob watched the men put the fire out.

"Ding-dong," went the bell again. The engine started back down the street.

Bob watched the engine go. Then he went back home and ate his dinner.

3b

Mary was going downtown to watch the parade. She skipped and ran along the street because she could hardly wait to get there. She was early and found a good place to stand.

Pretty soon she could hear the music of the bands coming down the main street. The men of the first band were dressed in scarlet, with white feathers in their hats. The men of the second band were clad in dark blue, with red feathers in their caps.

After them came the trucks loaded with flowers or fruit. Then there were cars filled with officers and their friends. Next came a company of soldiers in dark green uniforms. Last of all was another band dressed in white suits and yellow feathers.

4b

Mary's teacher took her class for a nature walk one sunshiny day last week. Every time the group came to a new plant, they would stop and examine it while the teacher explained its parts. She showed them how a bee gets its honey from flowers and how a bug had eaten part of the leaves from some plants. On a few plants, the flowers had fallen off, and seeds had begun to form.

Later, while they were looking at some blossoms, one boy spied a nest hidden in a tree. They were very quiet, hoping that the mother bird would return to feed her young ones. Sure enough, she quickly came back with a fat, juicy worm in her bill. She fed the young ones, chirped a little, and then flew away after more worms. Mary's teacher said that birds eat a great deal every day. They help us by eating insects that would destroy our plants and by eating weed seeds.

5b

Suppose that you have some beautiful poppies growing in your garden. Suppose, too, that you want to get some seed from them so that you can have more poppies like them next year. You must be sure, then, not to pick all the poppy flowers. If you do not leave some of the flowers on the plants, you will not have any seeds, for the flowers are the part of the plant that produces the seeds. There will not be any seeds if all the flowers are picked.

Most seeds come from flowers. The seeds of pine trees and of the trees and bushes of the pine family are formed in cones. But most other seeds come from flowers. More than 190,000 kinds of plants produce seeds, and all but about 700 produce their seeds in flowers.

Not all flowers are large and bright-colored like poppies. Probably you have seen many flowers that you did not know were flowers. Did you ever see any cottonwood flowers, or willow flowers, or grass flowers? Cottonwood trees and willow trees and grass have flowers, but their flowers are small and are not bright-colored. Many other plants have small flowers much like these.

6b

When the early settlers came to America, trade was carried on by barter or by using such things as tobacco, sugar, and furs as money. Sometimes the settlers used Indian wampum. Wampum was shells that were made into beads and was used by the Indians as decoration and as money. Of course, when more people came from Europe to settle in America, they found they would need money to pay workmen. A mason did not always want to take furs for his pay. A furrier did not always want his wages in grain or tools. People had to have coins, so they used whatever was available--English shillings, Swedish and Dutch money, and Spanish dollars, or "pieces of eight." The colonists soon found there were not enough of these to go around.

England would not let her Colonists make any money of their own. But in 1652, Massachusetts set up a mint and made her own coins anyway. Among these were the famous "pine-tree shillings." They were called this because the picture of a pine tree was stamped on them. These pine-tree shillings were made for thirty-four years, but they all had the same date on them. In this way the Colonists pretended that they were obeying England.

7b

One of the most beautiful and lasting kinds of building stone is marble. Marble may be pure white or colored, or it may have streaks in it. It can be polished so that it has a very smooth surface. All marble was once limestone. Limestone, deep under the ground, may be changed to marble by heat and pressure.

Granite is another very good building stone. It is formed from rock so hot that it is liquid. You may have seen pictures of liquid rock, or lava pouring out of volcanoes. Lava cools and becomes rock rather quickly. But granite is made from the hot liquid held underground. This rock cools very slowly. The liquid rock from which granite comes cools so slowly that the different materials in it separate from one another and form crystals. Granite is always a speckled rock because the different crystals in it are of different colors. Two minerals are always found in granite. They are quartz and feldspar. The dark speckles in granite are usually some other mineral. Granite makes very good building material because it is so hard. It can be beautifully polished, and the weather does not harm it.

8b

One very important reason for slow reading is lack of pre-planning. Many slow readers have not learned to adapt their rate to the difficulty of the material they are reading. They can see objects quickly with their eyes, as in looking through the window of a moving car, but when they read, the same quick movements are not present. They stop to look at each individual word as though that were necessary for gaining ideas. They tend to read their textbooks, magazines, and even the newspaper at about the same rate of speed.

It has been shown by many studies of good readers that the rate in fiction material should be two to three times as fast as that in non-fiction. Slow readers tend to "study" everything they read, while good readers vary their rate consciously according to the kinds of facts they want to get. For example, if a student is trying to find only one fact on a page, he certainly does not need to read the entire page. He can skim quickly over the page until he finds the fact he is seeking. If, on the other hand, he is expected to report critically on a piece of prose, he will need to read much more slowly and analytically.

2a

A hen had six little yellow chickens. One morning she took them for a walk. They looked for something to eat. They found some seeds and sand. A dog came to play with them. The hen did not like the dog. She flew at the dog and made him run away.

3a

Three boys built a house in the woods. They put a table and two old chairs in it. There was a basket full of apples under the table. One afternoon they went away and left the door open. When they came back, they found two little pigs eating the apples.

4a

A little girl got off the train all alone. There was nobody at the station to meet her. She asked the man inside the station where her mother was. He said that her mother could not get the car started. A man was trying to fix it. The little girl sat down to wait. A few minutes later a big car came around the corner with her mother in it. The little girl got in and they drove home.

5a

About one hundred and fifty years ago, in France, the first man went up in a balloon. His balloon was made of paper covered with strips of cloth to make it strong. A long rope kept it from going too high. Later this man took a friend up in the balloon with him. On this trip they rose over five hundred feet. The trip lasted thirty minutes. They came down several miles from where they started.

6a

Early settlers in America found that Indians would sell skins and land for glass beads. Many men earned their living by making glass beads and bottles. In 1827 a man invented a way to press molten glass into iron molds. The most famous glass works was in the town of Sandwich in Massachusetts. The Sandwich glass had a bright silvery appearance and it could be molded into very elaborate and attractive patterns. Beautiful lamps and candlesticks as well as all sorts of dishes were made from this glass. In many New England homes pieces of Sandwich glass are still found on display.

7a

Basketball is one of the more recent games. It was devised by a college instructor who desired a game to interpose between the football and baseball seasons. The game demands precision of movement, concentration, and great endurance. It is more popular in those localities where it does not compete with hockey. Opinion differs as to whether it is a satisfactory game for girls. It has been modified to make it less strenuous for them by restricting the playing area of each player. Some of the large Western universities have audiences of over twenty thousand at their conference games.

8a

Railroad communication developed rapidly just after the Civil War. Between 1865 and 1873, thirty-five thousand miles of track were laid. This doubled the distance people could travel by railroad. Some of the new roads connected important cities, and some extended westward beyond populated regions. Congress favored this sudden development by granting land to companies interested in furthering the expansion. Grants included territory lying within twenty miles of the proposed roadbed. Alternate sections were allotted to the railroad; those in between were reserved for homesteaders. The sale of sections of land owned by the railroad was made easier through this checkerboard arrangement.

2b

The policeman has a big brown horse named Pep. The children like to pat him. They give him apples and sugar to eat. Pep can do two tricks. He can shake hands. When the policeman says, "Count three," Pep paws the ground three times with his foot.

3b

One morning Jack woke up late. He had to dress in a hurry. He put on one of his black shoes, but he could not find the other. He could find only an old brown shoe. He put it on and ran to school. He got there just as the bell rang.

4b

Late in the summer a man started to build a house. He wanted to finish it before winter came. He had some men dig the cellar for him. Then he built the floor and the cellar steps himself. Before he could do anything more, he had to move away to another town. He left the house just as it was. He told the boys next door that they might use the cellar as a clubhouse.

5b

The first trains were pulled by horses. Later, engines were used to pull trains. The first engines could not go very fast and often broke down. Once there was a race between a train pulled by a horse and a train pulled by an engine. At first the horse was ahead because the engine had to start slowly. Then the engine passed the horse, but something in the engine broke and the horse won the race.

6b

Rubber came into general use about two hundred years ago. Columbus brought it back from his second voyage to America. It was named "rubber" because it was used to rub out pencil marks. Rubber comes mainly from Africa and Brazil. It is manufactured from the milky juice of the rubber tree. Rapid handling of the juice is necessary. It is collected early in the morning when the flow is greatest. At the factory, acid is added at once to make the rubber rise to the top. It is then treated in many different ways. Attempts to produce rubber in the United States have not been successful.

7b

Every four years, athletes from all over the world compete in a festival of sports. These great contests are called the Olympic Games, since they were first conducted at Mount Olympus by the ancient Greeks. They were part of a religious celebration in honor of Zeus, father of the Greek gods. They were discontinued long ago and were revived only in 1896 by men interested in continuing the custom. The first of the modern games were staged appropriately in Athens. Since that time they have been held in various countries. They stimulate world friendship and help to maintain interest in physical perfection.

8b

One of the most difficult political and economic problems of our Federal government is to prevent the development, throughout the country, of certain commercial interests at the expense of others. One very persistent attack has been against the development of large corporations. DeWitt in his study of the history of this struggle states that it has three distinct objectives. First, to find some satisfactory means to control and regulate the activities of large business corporations; second, to resist the tendency of corporations to exploit natural resources for their own benefit; third, to control tariffs which favor trusts and monopolies.

Appendix B: Student Protocol Notebook

This is an example of one student protocol notebook. It was used for directing the testing session and recording student responses. Each protocol notebook was generated separately for each student and each test session by a text processing system. Therefore, each student's protocol notebook contains the precise versions or alternate forms for him/her. The protocol notebook was essentially used as a script which was read and followed by the test examiner. The sample given here is for student id 117 who received test battery 601. Note that the notebook was changed slightly from pre- to post-test. Specifically, the nature of instructions and reminders to the test examiner were improved to guarantee a precise, standard use of the materials. The contents presented in this appendix are identical to that used in the study except that the format has been altered slightly.

1. Please record today's date and the time:
   Date:__________ Time:________

2. Please record your (the tester's) name:
   Name:________________

3. The following materials should be present for the testing session:

   a. The materials for the student to use:

      1) The 'Test Materials Notebook'

   b. The materials for the test administrator to use:

      1) The 'Test Protocol Notebook' for the particular student that is to be tested.

      2) An audio cassette taperecorder for recording the entire testing session.

      3) A 90 minute cassette tape.

      4) A directional microphone.

      5) A stop watch for timing various parts of the testing session.

4. The order of presentation of materials to the student will be determined by the order of materials in the 'Test Protocol Notebook'. This order should be:

   a. Standard Interview form

   b. Instant Word Recognition Measure: Using the Slosson Oral Reading Test and non standard procedures

   c. Decoded Word Recognition Measure: Using the Gates-McKillop Test

   d. Oral Reading Measure: Using the Gray Oral Reading Test

   e. Sullivan Oral Reading Placement Task

   f. Reading Comprehension Measure: Using Spache and Durrell Graded Selections and non standard procedures

   g. Listening Comprehension Measure: Using the Spache and Durrell Graded Selections and non standard procedures

   h. Word Comprehension Measure: Vocabulary from the Spache and Durrell Comprehension Selections

   i. Test Administrator Debriefing Form

5. Check to make sure that the student's name and code number are on the front cover of the 'Test Protocol Notebook' and that his/her code number is on all successive pages of the same document.

6. Avoid the use of the student's name during standard interview and oral reading tasks. Do not write the student's name on any of the materials.

7. DO NOT GIVE THE STUDENT HIS/HER TEST MATERIALS AT THIS TIME.

8. Check out the audio tape recorder before the student arrives. Actually record something and play it back to check volume levels, proper functioning of tape and recorder and microphone, etc. DO THIS EVERY TIME!!!!!!

Introduction to Student

After greeting the student and introducing yourself, orient the student to the session. If this is the first encounter with the student this year, use the first set of statements as your guide. If the student has already been diagnosed once this year, use the second set. Informally, and in your own words, orient the student using the appropriate guidelines.

1. STATEMENT SET 1: FOR THE FIRST DIAGNOSIS OF A STUDENT

   a. that you know the student has had problems with reading

   b. that the purpose of the session is to gather information so that you can figure out how to make the student a better reader

   c. that the session will last at least an hour. During most of that time the student will be reading either out loud or to him/herself.

   d. that during the session you will be jotting things down, using the stopwatch from time to time and taperecording the whole session.

   e. let the student ask you any questions s/he might have.

2. STATEMENT SET 2: FOR SUCCESSIVE DIAGNOSES OF A STUDENT

   a. that because of the way the reading program is setup here, since the student was tested at the beginning of the school year, we have to test him/her again.

   b. that, again, the session will last at least an hour. During most of that time the student will be reading either out loud or to him/herself.

   c. that, again, during the session you will be jotting things down, using the stopwatch from time to time and taperecording the whole session.

   d. let the student ask you any questions s/he might have.

<u>Standard Interview Form</u>

1. Please record the time. Time:________

2. TURN ON THE TAPE RECORDER!!! to record the entire testing session.

3. Read the following to the student:

   "I'M GOING TO ASK YOU SOME QUESTIONS. YOUR ANSWERS WILL HELP US FIGURE OUT THE BEST WAY TO HELP YOU BECOME A BETTER READER."

4. If this is the first encounter with the student this year, ask the first set of questions. If the student has already been diagnosed once this year, ask the second set of questions. DO NOT write down the student responses. The answers will be transcribed from the audio tape at a later date.

5. QUESTION SET 1: FOR THE FIRST DIAGNOSIS OF A STUDENT

   a. Why do you think you're having trouble reading?

   b. What kinds of books are easy to read for you?

   c. Do you remember when you started having trouble with reading?

      1) If "Yes", then, ask "When?"

      2) If "No", then, ask the student to think back to the fourth grade. Was everything alright then? Repeat until approximate grade established.

   d. When reading your school books, do you know most of the words right away?

   e. What do you do when you come to a word you don't know?

   f. When you listen to people read out loud, how can you tell if they're doing a good job?

   g. Are you having any trouble finishing your reading assignments in your classes?

   h. What do you read outside of school?

   i. When summarizing a fiction story, what kind of things should you include in the summary?

   j. What kind of reading is hard for you to summarize?

   k. Do your eyes bother you after you've been reading for a while?

l. Do you have any trouble seeing things clearly?

m. Do you have glasses? Should you wear them when reading? Do you?

n. Do you have any trouble hearing people when they talk to you?

o. What is punctuation for?

p. Have you been in any special reading classes before?

q. What kinds of things have you done in your reading classes?

r. Is there anything particular that you want to be able to do in reading by the end of this year?

6. QUESTION SET 2: FOR SUCCESSIVE DIAGNOSES OF A STUDENT

a. When reading your school books, do you know most of the words right away?

b. What do you do when you come to a word you don't know?

c. When you listen to people read out loud, how can you tell if they're doing a good job?

d. Are you having any trouble finishing your reading assignments in your classes?

e. When summarizing a fiction story, what kind of things should you include in the summary?

f. What kind of reading is hard for you to summarize?

g. Do your eyes bother you after you've been reading for a while?

h. Do you have any trouble seeing things clearly?

i. Do you have glasses? Should you wear them when reading? Do you?

j. Do you have any trouble hearing people when they talk to you?

k. What is punctuation for?

****
Affect rating during this task:

1.Did the student attend closely to the task?

YES   NO

2.Other comments on affect during this task?

Instant Word Recognition Measure

1. Please record the time. Time:________

2. Give the student his/her "Test Materials Booklet." Determine the appropriate form of the SORT for this student by looking at the next page of this protocol notebook. Then, turn to the proper page of the "Test Materials Notebook" for the student as indicated in its table of contents.

   ****THIS STUDENT WILL USE FORM SORT RA****

3. Read the following directions to the child:

   "HERE ARE SOME LISTS OF WORDS. I'LL ASK YOU TO READ SOME OF THESE LISTS. IF YOU DON'T KNOW A WORD RIGHT AWAY, TRY AND SOUND IT OUT. WE'RE GOING TO START WITH THE VERY FIRST LIST."

4. For each word, record student performance as follows:

   a. If the student reads the word with no hesitation, check the 'Instant' column.

   b. If the student correctly called the word within a slow count of five, check the 'Mediated' column.

   c. If the student does not correctly call the word within a slow count of five, write down verbal responses in the 'Miscall' column. An incorrectly accented word is considered a miscall. Write N.R. (No Response) if no response was made.

   d. After the five count, ask the student to continue with the next word.

5. Let the student proceed through the lists, in order, until s/he can only read 5 or fewer words correctly. A word is considered correct whether called instantly or mediated.

6. Close the student's "Test Materials Notebook." For each of the 20 words in the last list the student attempted, ask the following question: 'What does ________ mean?' Check the appropriate column: 'Correct Meaning' 'Incorrect Meaning' or 'No Response'.

REMEMBER: THE CRITERION IS 5 OR FEWER WORDS CORRECT - INSTANT OR MEDIATED.

Word	Ins.	Med.	Miscall	Cor.	Inc.	NR
1. jump						
2. want						
3. look						
4. play						
5. come						
6. is						
7. can						
8. one						
9. see						
10. ball						
11. little						
12. three						
13. baby						
14. run						
15. down						
16. mother						
17. here						
18. up						
19. help						
20. make						
TOTALS						

Word	Ins.	Med.	Miscall	Cor.	Inc.	NR
21. very						
22. under						
23. hill						
24. road						
25. dark						
26. horse						
27. came						
28. puppy						
29. was						
30. bump						
31. with						
32. ride						
33. along						
34. friends						
35. first						
36. food						
37. wish						
38. what						
39. basket						
40. live						
TOTALS						

Word	Ins.	Med.	Miscall	Cor.	Inc.	NR
41. better	___	___	___	___	___	___
42. hide	___	___	___	___	___	___
43. large	___	___	___	___	___	___
44. field	___	___	___	___	___	___
45. around	___	___	___	___	___	___
46. suddenly	___	___	___	___	___	___
47. forest	___	___	___	___	___	___
48. river	___	___	___	___	___	___
49. sheep	___	___	___	___	___	___
50. station	___	___	___	___	___	___
51. breakfast	___	___	___	___	___	___
52. across	___	___	___	___	___	___
53. heavy	___	___	___	___	___	___
54. lunch	___	___	___	___	___	___
55. happen	___	___	___	___	___	___
56. grass	___	___	___	___	___	___
57. game	___	___	___	___	___	___
58. farmer	___	___	___	___	___	___
59. stars	___	___	___	___	___	___
60. hope	___	___	___	___	___	___
TOTALS	___	___	___	___	___	___

Word	Ins.	Med.	Miscall	Cor.	Inc.	NR
61. hunger	___	___	___	___	___	___
62. evening	___	___	___	___	___	___
63. bench	___	___	___	___	___	___
64. excuse	___	___	___	___	___	___
65. grove	___	___	___	___	___	___
66. desire	___	___	___	___	___	___
67. perform	___	___	___	___	___	___
68. destroy	___	___	___	___	___	___
69. delicious	___	___	___	___	___	___
70. understood	___	___	___	___	___	___
71. safe	___	___	___	___	___	___
72. against	___	___	___	___	___	___
73. timid	___	___	___	___	___	___
74. ocean	___	___	___	___	___	___
75. damp	___	___	___	___	___	___
76. stream	___	___	___	___	___	___
77. empty	___	___	___	___	___	___
78. smash	___	___	___	___	___	___
79. stone	___	___	___	___	___	___
80. reward	___	___	___	___	___	___
TOTALS	___	___	___	___	___	___

Word	Ins.	Med.	Miscall	Cor.	Inc.	NR
81. claimed						
82. forehead						
83. serious						
84. speechless						
85. distant						
86. dainty						
87. anger						
88. courage						
89. slumber						
90. common						
91. harness						
92. flakes						
93. develop						
94. vacant						
95. price						
96. future						
97. appearance						
98. promptly						
99. region						
100. silence						
TOTALS						

Word	Ins.	Med.	Miscall	Cor.	Inc.	NR
101. dignity	___	___	___	___	___	___
102. generally	___	___	___	___	___	___
103. gracious	___	___	___	___	___	___
104. applause	___	___	___	___	___	___
105. define	___	___	___	___	___	___
106. cushion	___	___	___	___	___	___
107. jungle	___	___	___	___	___	___
108. haze	___	___	___	___	___	___
109. profitable	___	___	___	___	___	___
110. tailor	___	___	___	___	___	___
111. fragrant	___	___	___	___	___	___
112. obedient	___	___	___	___	___	___
113. marriage	___	___	___	___	___	___
114. presence	___	___	___	___	___	___
115. terrace	___	___	___	___	___	___
116. interfere	___	___	___	___	___	___
117. ambition	___	___	___	___	___	___
118. merchant	___	___	___	___	___	___
119. custom	___	___	___	___	___	___
120. extended	___	___	___	___	___	___
TOTALS	___	___	___	___	___	___

Word	Ins.	Med.	Miscall	Cor.	Inc.	NR
121. tremendous						
122. dungeon						
123. customary						
124. abundant						
125. malicious						
126. imaginary						
127. excellence						
128. consequently						
129. installed						
130. responsible						
131. yearning						
132. liquid						
133. medicine						
134. spectacular						
135. rebellion						
136. compliments						
137. infected						
138. inventory						
139. importance						
140. detained						
TOTALS						

Word	Ins.	Med.	Miscall	Cor.	Inc.	NR
141. reminiscence						
142. architecture						
143. environment						
144. counterfeit						
145. pensive						
146. tremor						
147. intricate						
148. administer						
149. continuously						
150. attentively						
151. industrious						
152. contemporary						
153. approximate						
154. standardize						
155. complexion						
156. exhausted						
157. malignant						
158. compassionate						
159. crisis						
160. society						
TOTALS						

Word	Ins.	Med.	Miscall	Cor.	Inc.	NR
161. nucleus	___	___	___	___	___	___
162. proportional	___	___	___	___	___	___
163. articulate	___	___	___	___	___	___
164. remarkably	___	___	___	___	___	___
165. nonchalant	___	___	___	___	___	___
166. memorandum	___	___	___	___	___	___
167. deprecate	___	___	___	___	___	___
168. contrasting	___	___	___	___	___	___
169. whimsical	___	___	___	___	___	___
170. evident	___	___	___	___	___	___
171. prairies	___	___	___	___	___	___
172. twilight	___	___	___	___	___	___
173. intangible	___	___	___	___	___	___
174. supplement	___	___	___	___	___	___
175. formulated	___	___	___	___	___	___
176. exuberant	___	___	___	___	___	___
177. antique	___	___	___	___	___	___
178. inducement	___	___	___	___	___	___
179. irrelevance	___	___	___	___	___	___
180. grotesque	___	___	___	___	___	___
TOTALS	___	___	___	___	___	___

Word	Ins.	Med.	Miscall	Cor.	Inc.	NR
181. sojourn	___	___	___	___	___	___
182. affable	___	___	___	___	___	___
183. auspicious	___	___	___	___	___	___
184. enamoured	___	___	___	___	___	___
185. futility	___	___	___	___	___	___
186. primordial	___	___	___	___	___	___
187. inadequacy	___	___	___	___	___	___
188. chastisement	___	___	___	___	___	___
189. envisage	___	___	___	___	___	___
190. pandemonium	___	___	___	___	___	___
191. panorama	___	___	___	___	___	___
192. compressible	___	___	___	___	___	___
193. gustatory	___	___	___	___	___	___
194. traverse	___	___	___	___	___	___
195. simultaneous	___	___	___	___	___	___
196. excruciating	___	___	___	___	___	___
197. contraband	___	___	___	___	___	___
198. decipher	___	___	___	___	___	___
199. scrupulous	___	___	___	___	___	___
200. facsimile	___	___	___	___	___	___
TOTALS	___	___	___	___	___	___

REMEMBER:

1. Close the student's test booklet.

2. Ask the student for the meaning of all words in the last list attempted.

3. Don't let the student spend more than a slow count of five on any one definition.

4. Do not indicate to the student that these words are from the lists just read.

5. If the student responds with a definition of a homonym, present the word again, spell it, and ask for the definition.

****
Affect rating during this task:

1.Did the student attend closely to the task?

YES NO

2.Other comments on affect during this task?

Decoded Word Recognition Measure

1. Did you give the vocabulary test on the SORT just completed? If not, DO IT NOW!

2. Please record the time. Time:________

3. Reopen the student test booklet to the appropriate form of the Gates-McKillop for this student. USE THE TABLE OF CONTENTS.

   ****THIS STUDENT WILL USE FORM GMK RA****

4. Read the following instructions to the student:

   "HERE ARE SOME LISTS OF WORD PARTS AND NONSENSE WORDS. PLEASE READ DOWN EACH LIST FOR ME."

5. As the student reads, record all miscalled word segments and nonsense words by writing down the inappropriate response in the space provided.

1.sk______
2.fr______
3.st______
4.sl______
5.ch______
6.cr______
7.gl______
8.gr______
9.sm______
10.dw______
11.pr______
12.cl______
13.wr______
14.th______
15.tw______
16.sw______
17.wh______
18.bl______
19.sh______
20.dr______
21.fl______
22.tr______
23.pl______

TOTALS_____

24.idge______
25.er______
26.ate______
27.ite______
28.ind______
29.een______
30.ible______
31.ock______
32.ome______
33.ell______
34.emp______
35.ickle______
36.ed______
37.ew______
38.ick______
39.ark______
40.asp______
41.oy______
42.ight______
43.able______
44.ing______
45.ow______
46.ack______

TOTALS_____

47.stind______
48.smow______
49.chible______
50.whickle______
51.proy______
52.dwer______
53.twasp______
54.shable______
55.flark______
56.thate______
57.slome______
58.trock______
59.swite______
60.gright______
61.plew______
62.glemp______
63.clidge______
64.bleen______
65.drack______
66.sked______
67.wrick______
68.fring______
69.crell______

TOTALS_____

****
Affect rating during this task:

1.Did the student attend closely to the task?

YES NO

2.Other comments on affect during this task?

<u>Gray Oral Reading Test</u>

1. Please record the time. Time:________

2. Now turn to the appropriate form of the Gray Oral Reading for this student. USE THE TABLE OF CONTENTS.

   ****THIS STUDENT WILL USE FORM GRAY RA****

3. Read the following instructions to the student:

   'YOU ARE GOING TO BE READING SOME PARAGRAPHS ALOUD. WE'LL READ THEM ONE AT A TIME. THEY GET HARDER AS YOU GO ALONG AND YOU WILL BE MAKING MISTAKES. DON'T WORRY ABOUT THAT. JUST DO THE BEST YOU CAN.'

4. Begin the oral reading with paragraph one. For each paragraph do the following:

   a. RESET THE STOPWATCH TO ZERO.

   b. Tell the student to begin. Then, you START THE STOPWATCH.

   c. Record the student's performance using the standard oral reading markings.

   d. When the paragraph has been completed stop the clock and record the time.

   e. Let the student continue through consecutive paragraphs until his rate exceeds two minutes for a given paragraph.

   f. Provide correct pronunciation (1) after 5 seconds or (2) on student request. Offer no other prompting.

REMEMBER: THE CRITERION IS TWO MINUTES. PROVIDE PROMPTING AS DESCRIBED ABOVE.

A3 START WATCH. HAVE STUDENT READ. RECORD ORAL READING. STOP WATCH. RECORD TIME. Time:________

One morning a boy made a boat. "Where can I play with it?" he asked. Father said, "Come with me in the car. We will take your boat with us." Soon the boy called, "Please stop. I see water. May I play here?" "Yes," said Father. "Have a good time."

A4 START WATCH. HAVE STUDENT READ. RECORD ORAL READING. STOP WATCH. RECORD TIME. Time:________

One day five children went out to play in the beautiful white snow. They played for a long time and then began to make snow animals.

One of the animals was a dog. Soon the dog next door came out of the house. When he saw the snow dog he said, "Bow-wow."

The children laughed. "Now we have a dog that can bark."

A5 START WATCH. HAVE STUDENT READ. RECORD ORAL READING. STOP WATCH. RECORD TIME. Time:________

It was pet day at the fair. The children were waiting for the parade of animals to begin. They had trained their pets to do many different tricks. Among them was a tall boy whose goat made trouble for him. It kicked and tried hard to break away. When it heard the band, it became quiet. During the parade it danced so well that it won a prize.

A6 START WATCH. HAVE STUDENT READ. RECORD ORAL READING. STOP WATCH. RECORD TIME. Time:________

Airplane pilots have many important jobs. They fly passengers, freight, and mail from one city to another. Sometimes they make dangerous rescues in land and sea accidents, and drop food where people or herds are starving. They bring strange animals from dense jungles to our zoos. They also serve as traffic police and spot speeding cars on highways.

A7 START WATCH. HAVE STUDENT READ. RECORD ORAL READING. STOP WATCH. RECORD TIME. Time:________

Hundreds of years ago, most of Europe was a very poor region. But China, a large country in eastern Asia, had many of the comforts of a rich, civilized nation. Only a few people from Europe had visited this distant region. One was the famous Marco Polo. He learned some of the languages that were spoken in China and served its great ruler for many years.

A8 START WATCH. HAVE STUDENT READ. RECORD ORAL READING. STOP WATCH. RECORD TIME. Time:________

The eager spectators who had cheered the plucky Warriors through eight hard-fought innings were silent. Only a run was required to defeat the much feared champions, who had previously defeated all opponents. The spectators had earlier criticized the umpire severely. Now their faces were tense with excitement as the players took their positions.

A9 START WATCH. HAVE STUDENT READ. RECORD ORAL READING. STOP WATCH. RECORD TIME. Time:________

The oil industry has been greatly increased by recent advances in science. Geologists have discovered new ways of locating veins of oil-producing rock. Problems of gusher control have been solved. Very effective also are newer methods of refining crude oil which have resulted in a higher ratio of quality fuel oil from a given volume of crude oil.

A10 START WATCH. HAVE STUDENT READ. RECORD ORAL READING. STOP WATCH. RECORD TIME. Time:________

In response to the impulse of habit Joseph rose and spoke as in former days. He spoke vigorously, continuously, and persuasively while the others listened attentively but in grim and contemptuous silence. Finally exhausted, Joseph hesitated for a moment; as often happens in such circumstances he became confused and was unable to resume speaking.

A11 START WATCH. HAVE STUDENT READ. RECORD ORAL READING. STOP WATCH. RECORD TIME. Time:________

Many of the hypotheses about physical phenomena formulated by early philosophers were inconsistent and in most cases could not be universally applied. In order to develop accurate principles very capable physicists, mathematicians, and statisticians had to cooperate wholeheartedly over long periods of time to verify numerous basic facts and assumptions.

A12 START WATCH. HAVE STUDENT READ. RECORD ORAL READING. STOP WATCH. RECORD TIME. Time:________

In a concluding lecture on sidereal spaces, the astronomer contrasted the infinitesimal difference in the distance of the moon from the earth at apogee and at perigee with the great difference in the distance of the earth from the sun at aphelion and at perihelion. The students interrogated him, evidencing precociousness and lucidity in expression.

A13 START WATCH. HAVE STUDENT READ. RECORD ORAL READING. STOP WATCH. RECORD TIME. Time:________

During a hiatus in the desultory firing, the apt lieutenant clambered wearily over the detritus piled against the redoubts. Beneath a canopy of empyrean blue lay the quiet, bucolic landscape, its pristine beauty now defiled by myriad diminutive promontories thrown up by the mortar shells, but radiating momentarily an inexplicable if spurious calm and peace.

****
Affect rating during this task:

1.Did the student attend closely to the task?

YES NO

2.Other comments on affect during this task?

1. Did the student generally use proper inflection while reading the passages? (circle one) YES NO

2. Was the student reasonably fluent? (circle one) YES NO

Sullivan Oral Reading Placement Task

1. Please record the time. Time:________

2. Turn to the Sullivan Reading Paragraphs. USE THE TABLE OF CONTENTS.

3. Read the following instructions to the student:

   "YOU ARE GOING TO BE READING SOME SHORT PARAGRAPHS ALOUD. PLEASE START WITH PARAGRAPH NUMBER ONE AND CONTINUE UNTIL I TELL YOU TO STOP."

4. Record the student's performance using the standard oral reading markings. Let the student continue through consecutive paragraphs until s/he makes three or more uncorrected miscalls in a paragraph.

5. Offer no prompts or other assistance as the student reads.

REMEMBER: THE CRITERION IS 3 OR MORE NONTRIVIAL UNCORRECTED MISCALLS. OFFER NO PROMPTING.

1. Sam is wet. A bat can fly. Ned sat up in bed. A fox bit Tim on the leg.

2. Bob and Pam sat on a raft. Bob had a fish in his hand. Pam had a ship.

3. Frank sent Nan a gift. She gives him a kiss to thank him.

4. A shark is a big fish that can harm a man. The man swims as fast as he can.

5. Bill has come in first. He has won. Jack is last. He has lost.

6. This bride has a long gown on. She stands in front of a church.

7. We can play lots of games with cards. These three cards are the ace of clubs, the king of clubs, and the queen of spades.

8. I eat three meals a day. For lunch I had a pound of ground meat, a slice of toast, and a cup of tea.

9. This nice young man has lost his sight. He is now blind, but he has a smart dog for a guide.

10. The player who throws the ball is the pitcher. The one who swings at it is the hitter. The man who crouches in back of the plate is the catcher.

11. Mister Smith is a father. His wife is a mother. They are both taller than their son or their daughter.

12. Jan White is washing her clothes in the washer and drying them in

the drier. When she is through, she will hang them up on the hangers.

13. The boy and the girl are seated at their desks. The boy is reading a book. The girl is writing a letter.

14. Winter is the coldest time of the year. The skies are gray. The days are cloudy, windy, and snowy.

15. In the brightly lit big tent, a monkey is riding on the back of a donkey. Each child has money in his hand. Billy has a quarter and Jimmy has a dollar bill.

16. The boy wearing the mask is trying to frighten the men and children. The puzzled poodle is barking at what he thinks is a witch.

17. This woman has a bonnet on her head and a bracelet on her left wrist. She is sitting at the window and feeding pretzels to a sparrow, a pigeon, and a squirrel.

18. In backing her car out of the garage, a woman driver hit a hydrant. Her luggage spilled out of the trunk in front of a cyclist who stopped all the traffic on the street.

19. The youthful rider is mounted on a prancing stallion. In the flat prairie below, a train is just pulling out of the station. To the far right, storm clouds are forming over the tall mountain peaks.

20. In the foreground, a barefoot woman in a bathrobe is using clothespins to pin up a bedspread. To her left, her younger brother is beating the dust out of a mattress. In the background, her uncle

is painting a bookcase.

****

Affect rating during this task:

1.Did the student attend closely to the task?

YES NO

2.Other comments on affect during this task?

1. Did the student generally use proper inflection while reading the passages? (circle one) YES NO

2. Was the student reasonably fluent? (circle one) YES NO

1. Please record the time. Time:________

2. Take a five minute break if the student wishes.

3. Turn the tape recorder off.

4. After the break, TURN THE TAPE OVER, turn the tape recorder back on and continue.

Silent Reading Measure

1. IS THE TAPE TURNED OVER? IS IT TURNED ON? IF NOT, DO SO.

2. Please record the time. Time:________

3. Turn to the appropriate form of the Silent Reading Measure for this student. USE THE TABLE OF CONTENTS.

   ****THIS STUDENT WILL USE FORM SPACHE RA****

4. Read the following instructions to the student:

   "YOU ARE GOING TO BE READING SOME SELECTIONS SILENTLY ONE AT A TIME. I'LL BE ASKING YOU ABOUT EACH SELECTION AFTER YOU READ IT. I'LL TELL YOU WHEN TO START. PLEASE TELL ME WHEN YOU HAVE FINISHED READING."

5. For each selection do the following:

   a. RESET THE STOPWATCH TO ZERO.

   b. Tell the student to begin. Then, you START THE STOPWATCH.

   c. When the student has finished the selection, stop the clock and record the time and cover the selection with a card.

6. For each selection conduct the comprehension procedure as follows:

   a. Ask the student, "TELL ME IN ONE SENTENCE WHAT THE MAIN IDEA IN THIS SELECTION IS." Check all responses made. If 'Other' write down response. Note: If student attempts to recite the whole selection, reprompt: "Please tell me just the MAIN idea."

   b. Ask any questions which follow and check all responses made. If 'Other' write down response.

   c. Say to the student, "TELL ME WHAT WAS IN THE SELECTION, STARTING AT THE BEGINNING." In the first column, number all memories in order of recall. When the student has finished, say, "CAN YOU REMEMBER ANYTHING ELSE?" Put a check in the second column next to the appropriate memory. Write any memories mentioned by the child which are not listed.

   d. Continue through all selections until the student takes more than 70 seconds or has fewer than 3 sequential memories. If, on the last selection, the student had fewer than 3 memories, then YOU reread the paragraph to him and reask the questions.

REMEMBER: THE CRITERION IS MORE THAN 70 SECONDS OR FEWER THAN 3 RECALLS.

Spache: 2a

1. START WATCH. YOU READ (if listening comprehension) or HAVE STUDENT READ (if silent comprehension). STOP WATCH. COVER SELECTION (if silent comprehension). RECORD TIME.
   Time:________

2. Tell me in one sentence what the main idea in this selection is.
   a. ____riding
   b. ____a wagon
   c. ____up and down hills
   d. ____(a boy; Bob)
   e. ____(a dog)
   f. ____other____________________

3. Do you think the dog enjoyed riding downhill in the wagon? Why?
   a. ____no
   b. ____he jumped out
   c. ____he was frightened
   d. ____he went too fast
   e. ____he thought he might get hurt
   f. ____other____________________
   g. ____yes
   h. ____other____________________

4. Tell me what the story was about, starting at the beginning.

__|__Bob has
__|__a little red wagon.
__|__He likes
__|__to ride in it.
__|__He pulls it
__|__slowly
__|__up the hill.
__|__Then he rides it
__|__quickly
__|__down again.
__|__One day
__|__he took his dog
__|__with him.
__|__He pulled
__|__the dog
__|__up the hill.
__|__Then they rode
__|__down the hill
__|__but the dog
__|__did not like
__|__to ride down.
__|__He jumped out
__|__of the red wagon.
__|__Bob went down
__|__by himself.
__|__Now he does not try
__|__to take his dog
__|__in the wagon.

NOW ASK: CAN YOU REMEMBER ANYTHING ELSE?

Spache: 3a

1. START WATCH. YOU READ (if listening comprehension) or HAVE STUDENT READ (if silent comprehension). STOP WATCH. COVER SELECTION (if silent comprehension). RECORD TIME.
   Time:________

2. Tell me in one sentence what the main idea in this selection is.
   a. ____Bob; a boy
   b. ____and Spotty; a dog
   c. ____take trips
   d. ____in the woods
   e. ____other____________________

3. Why did the dog have to sleep outside?
   a. ____he smelled bad
   b. ____he was being punished
   c. ____the skunk sprayed him
   d. ____he caught a skunk (half credit)
   e. ____other____________________

4. Tell me what the story was about, starting at the beginning.

__|__Bob has
__|__a brown
__|__and white dog
__|__named Spotty.
__|__He is called Spotty
__|__because he has brown
__|__spots
__|__on his nose.
__|__Bob always takes
__|__his dog
__|__on his trips
__|__to the woods.
__|__The dog
__|__helps scare
__|__the rabbits.
__|__Bob walks slowly,
__|__but his dog
__|__scampers through
__|__the leaves.
__|__One day
__|__Spotty
__|__left Bob
__|__and went off
__|__by himself.
__|__Bob called
__|__and whistled,
__|__but the dog
__|__did not come back
__|__to him.
__|__After a while
__|__Bob heard
__|__the dog
__|__barking
__|__a long way off.
__|__Bob walked
__|__toward the sound
__|__of the barking
__|__until he found
__|__the dog.
__|__Spotty thought
__|__he had caught
__|__a black and white
__|__kiten
__|__But it wasn't a kitten.
__|__it was a skunk.
__|__That night
__|__the dog
__|__had to sleep
__|__outside.

NOW ASK: CAN YOU REMEMBER ANYTHING ELSE?

Spache: 4a

1. START WATCH. YOU READ (if listening comprehension) or HAVE STUDENT READ (if silent comprehension). STOP WATCH. COVER SELECTION (if silent comprehension). RECORD TIME.
   Time:________

2. Tell me in one sentence what the main idea in this selection is.
   a. ____Bob's; a boy's
   b. ____trip
   c. ____to the city market; to the fruit market
   d. ____(fruit)
   e. ____other____________________

3. Why did Bob have trouble buying fruit in a place that was filled with fruits and vegetables?
   a. ____they sell only in large quantities (by the crate)
   b. ____Bob didn't want to buy enough
   c. ____other____________________

4. Tell me what the story was about, starting at the beginning.

__|__Yesterday
__|__Bob
__|__took a trip
__|__to a city market
__|__that was somewhat like
__|__a store
__|__but a great deal bigger.
__|__It didn't have
__|__any bread
__|__or canned goods
__|__like the grocery stores.
__|__But there were a great many big boxes
__|__of vegetables
__|__and fruits.
__|__Bob was hungry
__|__and wanted just one plum
__|__or cherry
__|__to taste.
__|__He wondered
__|__if one of the men
__|__would sell him
__|__just one plum.
__|__Everyone
__|__was buying the fruit
__|__and vegetables
__|__by the whole crate.
__|__When Bob
__|__asked the man
__|__to sell him
__|__one plum
__|__he laughed
__|__and gave Bob
__|__an extra large plum
__|__wrapped in paper
__|__but wouldn't take
__|__any money.
__|__As he walked along
__|__eating the plum
__|__Bob watched
__|__the men
__|__unloading the trucks
__|__and big trailers.
__|__They would chop open
__|__the top
__|__of the crate
__|__so that anyone
__|__could see
__|__the fruit.
__|__If a buyer
__|__liked the fruit
__|__and was willing to pay
__|__the price

__|__he might buy
__|__the entire truckload.

NOW ASK: CAN YOU REMEMBER ANYTHING ELSE?

Spache: 5a

1. START WATCH. YOU READ (if listening comprehension) or HAVE STUDENT READ (if silent comprehension). STOP WATCH. COVER SELECTION (if silent comprehension). RECORD TIME.
   Time:________

2. Tell me in one sentence what the main idea in this selection is.
   a. ____ship's boy; a boy; John Paul
   b. ____learns to be a sailor
   c. ____loves to sail
   d. ____(to America)
   e. ____other____________________

3. What leads you to believe that John Paul sailed on many voyages?
   a. ____he visited America often
   b. ____he learned to be an expert sailor
   c. ____he could sail with the best of men
   d. ____other____________________

4. Tell me what the story was about, starting at the beginning.

__|__As a ship's boy,
__|__John Paul
__|__had all sorts of odd jobs
__|__on board.
__|__Sometimes he scrubbed decks
__|__or helped the cook.
__|__He cleaned the captain's cabin
__|__and ran errands
__|__but he had other duties
__|__that pleased him more.
__|__He helped to clean the guns
__|__which the merchant ship
__|__carried for protection.
__|__Several times
__|__he stood behind the big wheel
__|__to steer the ship.
__|__Captain Benson
__|__wrote in the ship's log or daily record
__|__that the trip was calm
__|__and smooth-sailing.
__|__Nothing unusual happened
__|__but every day was a real adventure
__|__for the new ship's boy.
__|__At the end of the voyage
__|__it was a thrill
__|__to sight land.
__|__When the ship docked
__|__near Fredericksburg, Virginia,
__|__John Paul was waiting
__|__to go on shore.
__|__John Paul's brother
__|__had a tailor shop
__|__in Fredericksburg
__|__and was very happy
__|__in his new home.
__|__He was eager to talk
__|__about the wonderful country
__|__but John Paul already loved America.
__|__During the next few years
__|__John Paul visited America often.
__|__He became used to the free
__|__and democratic ways
__|__of the new country.
__|__Meanwhile he had learned
__|__to be an expert sailor.
__|__Although he was not tall
__|__he was strong
__|__and quick.
__|__With his long arms
__|__he could haul or trim a sail
__|__with the best of men.

NOW ASK: CAN YOU REMEMBER ANYTHING ELSE?

Spache: 6a

1. START WATCH. YOU READ (if listening comprehension) or HAVE STUDENT READ (if silent comprehension). STOP WATCH. COVER SELECTION (if silent comprehension). RECORD TIME.
   Time:________

2. Tell me in one sentence what the main idea in this selection is.
   a. ____description
   b. ____of mastodons; of ancient elephants
   c. ____(elephants)
   d. ____other____________________

3. Why do you suppose that the elephants of today do not have woolly hair?
   a. ____they live in warm regions
   b. ____they don't need it
   c. ____they are not the same kind of elephant
   d. ____they are a different kind
   e. ____other____________________

4. Tell me what the story was about, starting at the beginning.

__|__Elephants
__|__are found wild today
__|__only in warm regions
__|__in tropical Africa
__|__and in India.
__|__The story was very different 50 thousand years ago.
__|__Then, two species of the elephant family
__|__roamed North America
__|__and Europe
__|__in vast numbers.
__|__One of them was the mastodon.
__|__The mastodon lived in the eastern part of our country
__|__during the peiod of the Great Ice Age.
__|__In the swamps
__|__that were formed when the ice disappeared,
__|__many of the huge creatures
__|__were trapped and killed.
__|__We have found
__|__some of their skeletons.
__|__At a glance, the mastodon must have looked much like the elephants of today
__|__except that it was covered with coarse, woolly hair
__|__and its tusks
__|__were much larger.
__|__It was probably heavier than the elephants we know
__|__but not taller.
__|__Its head was flatter
__|__and its lower jaw longer.
__|__Its teeth were not like the teeth of the elephants of today.
__|__More than 200 years ago,
__|__the people of New England
__|__found bones
__|__of the mastodon
__|__when they dug ditches
__|__to drain swamps.
__|__At first they thought that the bones they found were bones of giant people.
__|__When they found teeth
__|__that weighed more than four pounds apiece
__|__they decided that the giants were giants indeed.

NOW ASK: CAN YOU REMEMBER ANYTHING ELSE?

Spache: 7a

1. START WATCH. YOU READ (if listening comprehension) or HAVE STUDENT READ (if silent comprehension). STOP WATCH. COVER SELECTION (if silent comprehension). RECORD TIME.
   Time:________

2. Tell me in one sentence what the main idea in this selection is.
   a. ____using different reading speeds
   b. ____for different reading materials
   c. ____for different purposes
   d. ____(reading)
   e. ____other____________________

3. Why do you think slow readers tend to study everything they read.
   a. ____they're afraid they will miss something important
   b. ____they don't realize that reading speed can be adjusted
   c. ____other____________________

4. Tell me what the story was about, starting at the beginning.

__|__Just as in driving a car,
__|__we use at least three speeds
__|__in reading.
__|__High gear in reading
__|__is called skimming,
__|__while studying
__|__is reading in low gear.
__|__Between these two, at a second gear
__|__is what might be called a moderate speed of reading.
__|__As you may have heard, the good reader
__|__adapts his rate
__|__to the purpose of his reading.
__|__The rate he uses
__|__is determined by how much he wants to get out of the material
__|__he is reading.
__|__His rate is also influenced
__|__by the difficulty
__|__of the reading material.
__|__Thus, he shifts
__|__from gear to gear
__|__according to the amount he wants to retain
__|__or how difficult he finds the going.
__|__Skimming is useful
__|__for a number of situations in reading.
__|__We can use it
__|__when looking for a particular fact on a page
__|__or in a table.
__|__It is also appropriate when we have to cover a large amount of material
__|__that is not too interesting
__|__or too important.
__|__Skimming may also be used to determine the general trends or ideas
__|__of a selection
__|__when we do not have to know the fine details.
__|__It is also helpful
__|__when we are making a quick brush-up
__|__before recitation.
__|__Finally, it is very useful
__|__as the speed at which we would do pre-reading
__|__before studying intensively.

NOW ASK: CAN YOU REMEMBER ANYTHING ELSE?

Spache: 8a

1. START WATCH. YOU READ (if listening comprehension) or HAVE STUDENT READ (if silent comprehension). STOP WATCH. COVER SELECTION (if silent comprehension). RECORD TIME.
   Time:________

2. Tell me in one sentence what is the main idea in this selection.
   a. ____Exploration
   b. ____by Lewis
   c. ____and Clark
   d. ____of the Northwest Territories
   e. ____other

3. Why did the expedition follow major rivers?
   a. ____provided water
   b. ____provided a pathway
   c. ____could use boats
   d. ____other

4. Tell me what the story was about, starting at the beginning.

__|__President Thomas Jefferson,
__|__in 1804
__|__commissioned an expedition
__|__to go into the Northwest Territory
__|__to explore the land
__|__that was bought
__|__in the now-famous Louisiana Purchase.
__|__Meriwether Lewis,
__|__private secretary to the President,
__|__and Captain William Clark
__|__of the United States Army
__|__headed the 26 men
__|__who started up the Missouri River
__|__from St. Louis
__|__on May 21, 1804.
__|__On July 18 of that year
__|__the group reached the southwest corner
__|__of the present state of Iowa,
__|__proceeded northward along the Missouri,
__|__and traversed parts of Iowa many times.
__|__The Lewis and Clark State Park
__|__west of Onawa
__|__was named in honor of these explorers.
__|__Sergeant Charles Floyd,
__|__a Kentucky backwoodsman
__|__and one of the most competent men of the party,
__|__became ill
__|__on August 19, 1804
__|__and died the following day.
__|__His body was laid to rest
__|__upon a high bluff
__|__near the present site of Sioux City
__|__where it is marked
__|__with a tall monument.
__|__The Floyd River
__|__and Sergeant Bluff
__|__were named in his honor.
__|__The expedition proceeded westward
__|__to the mouth of the Columbia River,
__|__and returned to Washington
__|__during the early months of 1807.
__|__Lewis
__|__was appointed governor
__|__of the Louisiana Territory
__|__and Clark
__|__was named governor
__|__of the Northwest Territory.

NOW ASK: CAN YOU REMEMBER ANYTHING ELSE?

****
Affect rating during this task:

1.Did the student attend closely to the task?

YES   NO

2.Other comments on affect during this task?

Listening Comprehension Measure

1. Please record the time. Time:________

2. Take the test booklet away from the student. Turn to the appropriate form of the Listening Comprehension Measure for this student. USE THE TABLE OF CONTENTS. You will be reading the selections to the student FROM the STUDENT'S test booklet.

   ****THIS STUDENT WILL USE FORM DURRELL RA****

3. Read the following instructions to the student:

   "YOU ARE GOING TO BE LISTENING TO SOME SELECTIONS. I'LL BE ASKING YOU ABOUT EACH SELECTION AFTER YOU'VE LISTENED TO IT.

4. Read at a moderate pace using normal inflection. Start the stopwatch when you begin reading and stop it when done. Record the time in the appropriate place for the selection.

5. For each selection conduct the comprehension procedure as follows:

   a. Ask the student, "TELL ME IN ONE SENTENCE WHAT THE MAIN IDEA IN THIS SELECTION IS." Check all responses made. If 'Other' write down response. Note: If student attempts to recite the whole selection, reprompt: "Please tell me just the MAIN idea."

   b. Ask any questions which follow and check all responses made. If 'Other' write down response.

   c. Say to the student, "TELL ME WHAT WAS IN THE SELECTION, STARTING AT THE BEGINNING." In the first column, number all memories in order of recall. When the student has finished say, "CAN YOU REMEMBER ANYTHING ELSE?" Put a check in the second column next to the appropriate memory. Write any memories mentioned by the child which are not listed.

   d. Continue through all selections until the student has fewer than 3 memories in sequential recall.

Durrell: 2a

1. START WATCH. YOU READ (if listening comprehension) or HAVE STUDENT READ (if silent comprehension). STOP WATCH. COVER SELECTION (if silent comprehension). RECORD TIME.
   Time:________

2. Tell me in one sentence what the main idea in this selection is.
   a. ____chickens; hen; chicks
   b. ____looking
   c. ____for food
   d. ____(dog)
   e. ____(bird)
   f. ____other____________________

3. Why did the hen fly at the dog?
   a. ____to frighten the dog away
   b. ____to protect the chicks
   c. ____because the dog bothered her
   d. ____the hen didn't like the dog
   e. ____to protect the food
   f. ____the dog was playing with her chickens
   g. ____other____________________

4. Tell me what the story was about, starting at the beginning.

__|__A hen had
__|__six little yellow chickens.
__|__One morning
__|__she took them for a walk.
__|__They looked for
__|__something to eat.
__|__They found some seeds and
__|__sand.
__|__A dog came
__|__to play with them.
__|__The hen
__|__did not like the dog.
__|__She flew at the dog
__|__and made him run away.

NOW ASK: CAN YOU REMEMBER ANYTHING ELSE?

Durrell: 3a

1. START WATCH. YOU READ (if listening comprehension) or HAVE STUDENT READ (if silent comprehension). STOP WATCH. COVER SELECTION (if silent comprehension). RECORD TIME.
   Time:________

2. Tell me in one sentence what the main idea in this selection is.
   a. ____boys
   b. ____built house
   c. ____in woods
   d. ____(pigs)
   e. ____(apples)
   f. ____other____________________

3. Why were the pigs able to get to the apples?
   a. ____door was open
   b. ____apples were on the floor
   c. ____the boys were not there
   d. ____other____________________

4. Tell me what the story was about, starting at the beginning.

__|__Three boys
__|__built a house
__|__in the woods.
__|__They put a table
__|__and two old chairs in it.
__|__There was a basket
__|__full of apples
__|__under the table.
__|__One afternoon
__|__they went away
__|__and left the door open.
__|__When they came back
__|__they found two little pigs
__|__eating the apples.

NOW ASK: CAN YOU REMEMBER ANYTHING ELSE?

Durrell: 4a

1. START WATCH. YOU READ (if listening comprehension) or HAVE STUDENT READ (if silent comprehension). STOP WATCH. COVER SELECTION (if silent comprehension). RECORD TIME.
   Time:________

2. Tell me in one sentence what the main idea in this selection is.
   a. ____little girl
   b. ____waiting; nobody there to pick her up
   c. ____at train station; after train ride; getting off the train
   d. ____for mother
   e. ____other____________________

3. Why did the girl have to wait at the station?
   a. ____her mother wasn't there
   b. ____her mother's car broke down
   c. ____no one was there
   d. ____other____________________

4. How did the man inside the station know why the girl's mother was late?
   a. ____the mother called the station
   b. ____the mother sent a message
   c. ____other____________________

5. Tell me what the story was about, starting at the beginning.

__|__A little girl
__|__got off the train
__|__all alone.
__|__There was nobody
__|__at the station
__|__to meet her.
__|__She asked the man
__|__inside the station
__|__where her mother was.
__|__He said that her mother
__|__could not get the car started.
__|__A man was trying to fix it.
__|__The little girl sat down
__|__to wait.
__|__A few minutes later
__|__a big car
__|__came around the corner
__|__with her mother in it.
__|__The little girl got in
__|__and they drove home.

NOW ASK: CAN YOU REMEMBER ANYTHING ELSE?

Durrell: 5a

1. START WATCH. YOU READ (if listening comprehension) or HAVE STUDENT READ (if silent comprehension). STOP WATCH. COVER SELECTION (if silent comprehension). RECORD TIME.
   Time:________

2. Tell me in one sentence what the main idea in this selection is.
   a. ____first man(note: the word FIRST must be used)
   b. ____went up
   c. ____in a balloon
   d. ____other____________________

3. On the second trip, why did they probably not use a long rope?
   a. ____balloon travelled too far
   b. ____wanted the balloon to go higher
   c. ____didn't have a long enough rope
   d. ____he knew how to handle the balloon
   e. ____he already knew it was safe
   f. ____so he could see how far and high they could go
   g. ____they didn't need it (half credit)
   h. ____other____________________

4. Tell me what the story was about, starting at the beginning.

__|__About one hundred and fifty
__|__years ago
__|__in France,
__|__the first man
__|__went up in a balloon.
__|__His balloon was made of
__|__paper
__|__covered with strips of cloth
__|__to make it strong.
__|__A long rope kept it
__|__from going too high.
__|__Later this man took a friend
__|__up in the balloon with him.
__|__On this trip they rose
__|__over five hundred feet.
__|__The trip lasted
__|__thirty minutes.
__|__They came down
__|__several miles
__|__from where they started.

NOW ASK: CAN YOU REMEMBER ANYTHING ELSE?

Durrell: 6a

1. START WATCH. YOU READ (if listening comprehension) or HAVE STUDENT READ (if silent comprehension). STOP WATCH. COVER SELECTION (if silent comprehension). RECORD TIME.
   Time:________

2. Tell me in one sentence what the main idea in this selection is.
   a. ____glass
   b. ____early American
   c. ____uses of
   d. ____Sandwich glass
   e. ____(glass beads)
   f. ____other____________________

3. How did the Sandwich glass works make its glass?
   a. ____the way the man did
   b. ____by using iron molds
   c. ____they melted it down
   d. ____other____________________

4. Tell me what the story was about, starting at the beginning.

__|__Early settlers
__|__in America
__|__found that Indians
__|__would sell skins and land
__|__for glass beads.
__|__Many men earned their living
__|__by making glass beads
__|__and bottles.
__|__In 1827
__|__a man invented a way
__|__to press molten glass
__|__into iron molds.
__|__The most famous glass works
__|__was in the town of Sandwich in Massachusetts.
__|__The Sandwich glass had
__|__a bright silvery appearance
__|__and it could be molded into
__|__very elaborate and attractive patterns.
__|__Beautiful lamps and candlesticks
__|__as well as all sorts of dishes
__|__were made from this glass.
__|__In many New England homes
__|__pieces of Sandwich glass
__|__are still found on display.

NOW ASK: CAN YOU REMEMBER ANYTHING ELSE?

Durrell: 7a

1. START WATCH. YOU READ (if listening comprehension) or HAVE STUDENT READ (if silent comprehension). STOP WATCH. COVER SELECTION (if silent comprehension). RECORD TIME.
   Time:________

2. Tell me in one sentence what the main idea in this selection is.
   a. ____basketball
   b. ____in college
   c. ____(invented)
   d. ____other____________________

3. Why was the playing area for each player restricted?
   a. ____to make it easier for girls
   b. ____they thought girls weren't as strong as boys
   c. ____other____________________

4. Tell me what the story was about, starting at the beginning.

__|__Basketball
__|__is one of the more recent games.
__|__It was devised
__|__by a college instructor
__|__who desired a game to interpose
__|__between the football
__|__and baseball seasons.
__|__The game demands
__|__precison of movement,
__|__concentration,
__|__and great endurance.
__|__It is more popular
__|__in those localities where
__|__it does not compete with hockey.
__|__Opinion differs as to whether
__|__it is a satisfactory game
__|__for girls.
__|__It has been modified
__|__to make it less strenuous
__|__for them
__|__by restricting the playing area
__|__of each player.
__|__Some of the large Western
__|__universities
__|__have audiences
__|__of over twenty thousand
__|__at their conference games.

NOW ASK: CAN YOU REMEMBER ANYTHING ELSE?

Durrell: 8a

1. START WATCH. YOU READ (if listening comprehension) or HAVE STUDENT READ (if silent comprehension). STOP WATCH. COVER SELECTION (if silent comprehension). RECORD TIME.
   Time:________

2. Tell me in one sentence what the main idea in this selection is.
   a. ____growth of
   b. ____railroads; trains; train companies
   c. ____after Civil War
   d. ____other____________________

3. Why do you think it would be important to extend the railroads beyond populated regions?
   a. ____people could settle new land
   b. ____resources could be transported back into the populated regions
   c. ____other____________________

4. How did Congress encourage the development of railroads?
   a. ____gave land
   b. ____other____________________

5. Tell me what the story was about, starting at the beginning.

__|__Railroad communication
__|__developed rapidly
__|__just after the Civil War.
__|__Between 1865 and 1873,
__|__thirty-five thousand
__|__miles of track were laid.
__|__This doubled the distance
__|__people could travel
__|__by railroad.
__|__Some of the new roads
__|__connected important cities
__|__and some extended westward beyond populated regions.
__|__Congress
__|__favored this sudden
__|__development
__|__by granting land to companies
__|__interested in furthering
__|__the expansion.
__|__Grants included territory
__|__lying within twenty miles
__|__of the proposed roadbed.
__|__Alternate sections
__|__were allotted to the railroad.
__|__Those in between were
__|__reserved for homesteaders.
__|__The sale of sections of land
__|__owned by the railroad
__|__was made easier
__|__through this checkerboard
__|__arrangement.

NOW ASK: CAN YOU REMEMBER ANYTHING ELSE?

****
Affect rating during this task:

1.Did the student attend closely to the task?

YES NO

2.Other comments on affect during this task?

Word Comprehension Measure

1. Please record the time. Time:________

2. For each of the words in the following two lists, ask the student the following question:

   "WHAT DOES ________ MEAN?"

3. For each word, check the appropriate column: 'Correct Meaning', 'Incorrect Meaning', or 'No Response'.

Vocabulary for SPACHE RA

Word	Cor.	Inc.	NR
1.scamper	___	___	___
2.crate	___	___	___
3.merchant ship	___	___	___
4.ship's log	___	___	___
5.species	___	___	___
6.coarse (spell it)	___	___	___
7.tusks	___	___	___
8.skimming	___	___	___
9.moderate	___	___	___
10.retain	___	___	___
11.recitation	___	___	___
12.intensively	___	___	___
13.expedition	___	___	___
14.competent	___	___	___

Vocabulary for DURRELL RA

Word	Cor.	Inc.	NR
1.elaborate	___	___	___
2.attractive	___	___	___
3.molten	___	___	___
4.devised	___	___	___
5.recent	___	___	___
6.precision	___	___	___
7.endurance	___	___	___
8.localities	___	___	___
9.modified	___	___	___
10.strenuous	___	___	___
11.restructuring	___	___	___
12.extended	___	___	___
13.allotted	___	___	___
14.expansion	___	___	___

****
Affect rating during this task:

1.Did the student attend closely to the task?

YES   NO

2.Other comments on affect during this task?

Test Administrator Debriefing Form

1. **
DO NOT DISMISS THE STUDENT YET.
****
Examine the entire test booklet to make sure that all tests are completed. Check them off.
____Interview
____SORT
____SORT Vocabulary
____Gates-Mckillop
____Gray
____Sullivan
____Silent Reading
____Listening Comprehension
____Vocabulary
If all are done, DISMISS the student.

2. Please record the time. Time:________
Please answer the following questions concerning the just completed testing session.

3. What was the student's overall affect?
____Animated
____Flat
____Sullen
____Hostile
____Passive
____Attentive
____Fidgety
____Nervous

4. What was the student's mean length of utterance?
____One word
____Phrase
____Sentence
____Multiple Sentences

5. What was the student's general sophistication of language?
____High
____Average
____Low

6. Did the student's speech patterns demonstrate any major differences from standard English?
____No
____Yes. Specify:________________

7. While reading silently and independently, what were the student's characteristics?
____Raced through material

____Squirmed
____Was attentive
____Gave up
____Asked for assistance
____Other. Specify:____________________

8. Did the student appear to have any difficulty hearing?
   ____No
   ____Yes. Specify:________________

9. Did the student appear to have any difficulty with the visual aspects of reading?
   ____Yes
   ____Rubbing eyes
   ____Watery eyes
   ____Keeping text close to eyes
   ____No

10. Enter any other comments about the session that have not been recorded elsewhere.

____

____

____

____

____

Appendix C: Sample Simulated Clinician Diagnostic Workup

Below is a sample run of the computer diagnostic aid. Note that the system assumes that the standard test battery has been given and asks for data based on that assumption. Clearly, in a production system, it would be more effective to use a micro computer or terminal connected to a mainframe computer to assist with test administration and automatically collect, integrate, and feedback the data.

```
USER:XSZ8 TASK: 2043
JUL 8, 1981 15:55:00
The Multiple Object Simulated Encounter System (MOSES)
Version: Experimental
Number: 1.02
Date: August 29, 1980
Developed by the Institute for Research on Teaching
Michigan State University
No system messages.
.
.
.
The Simulated Clinician based on the Model of Reading and Learning
is now in control. In which of the following contexts do you wish
the Simulated Clinician to operate?

1.Computer Aid in Diagnosis and Remediation
2.Computer Aid in Followup
3.Computer Directed Diagnosis and Remediation
4.Computer Directed Followup
5.Preceptor Instruction

Enter a number from 1 to 5 corresponding to your choice:
1

Please identify the clinician.
Enter the clinician's last name:
WAGNER
Enter the clinician's first name:
CHRIS
Enter the clinician's middle initial:
C
Enter the clinician's identification number (CIN):
1
Is the following data correct?
CLINICIAN LAST NAME :
WAGNER
CLINICIAN FIRST NAME :
CHRIS
CLINICIAN MIDDLE INITIAL :
C
CLINICIAN IDENTIFICATION NUMBER :
1
Y

Please identify the client.
Enter the client's last name:
SUBJECT 117#1
Enter the client's first name:
-
Enter the client's middle initial:
-
Enter the client's identification number (CIN):
```

0
Enter the client's grade placement as a number between 1 and 20.
7
Is the following data correct?
CLIENT LAST NAME :
SUBJECT 117#1
CLIENT FIRST NAME :
-
CLIENT MIDDLE INITIAL :
-
CLIENT IDENTIFICATION NUMBER :
0
Y
Certain routine information must be collected from the client.
The information collection procedures will require approximately
one hour of clinician time. The information collection procedures
are described both here in the computer files and the external
document entitled "Otto Junior High Initial Data Collection Procedures."

Please follow the procedures and collect the required information
from the client.

Do you need a description of any of the information collection
procedures?
N

Concerning the CLIPIR Standard Interview Procedure:

Did the reader's response to the question:
"When reading your school books, do you know most of the words right
away?"
indicate a potential problem with Instant Word Recognition ability?
N
Did the reader's response to the question:
"What do you do when you come to a word you don't know?"
indicate a potential problem with Decoded Word Recognition ability?
Y
Did the reader's response to the question:
"When you listen to people read out loud, how can you tell if they're
doing a good job?"
indiate a potential problem with oral reading?
Y
Did the reader's response to the questions:
"What kind of reading is hard for you to remember?"
or
"Do you have trouble remembering what you read?"
or
"What kind of books are hard to read?"
indiate a potential problem with silent reading?
N
Did the reader's response to the questions:
"Are you having any trouble finishing your reading assignments in
your classes?"
or

"Is there anything particular that you want to be able to do in reading by the end of this year?"
indiate a potential problem with motivation to read?
N
Did the reader's response to the questions:
"What is punctuation for?"
indicate a potential problem with punctuation usage?
Y
Did the reader's response to the questions:
"Do your eyes bother you after you've been reading for a while?"
and
"Do you have trouble seeing things clearly?
indiate a potential problem with visual acuity?
Y
Did the reader's response to the question:
"Do you have any trouble hearing people when they talk to you?
indicate a potential problem with auditory acuity?
Y
Did the reader's response to the question:
"What do you read outside class?"
or any other impressions from the session
indicate a potential lack of adequate reading practice?
N

Standard Interview Data Collection Complete.
OConcerning the CLIPIR Sort Procedure:
ODid the reader attempt list 1?
Y
How many words did the reader recognize INSTANTLY on list 1?
19
How many words did the reader recognize through MEDIATION on list 1?
0
How many words did the reader MISCALL on list 1?
1
ODid the reader attempt list 2?
Y
How many words did the reader recognize INSTANTLY on list 2?
15
How many words did the reader recognize through MEDIATION on list 2?
1
How many words did the reader MISCALL on list 2?
4
ODid the reader attempt list 3?
Y
How many words did the reader recognize INSTANTLY on list 3?
11
How many words did the reader recognize through MEDIATION on list 3?
3
How many words did the reader MISCALL on list 3?
6
ODid the reader attempt list 4?
Y
How many words did the reader recognize INSTANTLY on list 4?
6

How many words did the reader recognize through MEDIATION on list 4?
2
How many words did the reader MISCALL on list 4?
12
ODid the reader attempt list 5?
Y
How many words did the reader recognize INSTANTLY on list 5?
4
How many words did the reader recognize through MEDIATION on list 5?
2
How many words did the reader MISCALL on list 5?
14
ODid the reader attempt list 6?
N
ONow, concerning the vocabulary measure on the last list.
For list 5, how many words were correctly defined?
1
For list 5, how many words were incorrectly defined?
0
For list 5, how many definitions were not attempted?
6
Review the pattern of miscalls on the lists.
On how many of the miscalls were only the initial letters correct?
9
On how many of the miscalls were only the final letters correct?
0
For how many of the miscalls was the miscall a different word that was shaped like the stimulus word?
21
For how many of the miscalls was there confusion of similar letters?
0
OConcerning the CLIPIR Gates-Mckillop Procedure
How many of the consonant clusters were called correctly?
13
How many of the phonograms were called correctly?
13
How many of the nonsense words were called correctly?
3
Review the miscues. Did the miscues indicate any confusion of similar letters?
N
OConcerning the CLIPIR Gray Oral Reading Procedure:
Did the reader attempt selection 3
Y
How many seconds did it take the reader to read selection 1
35
How many errors did the reader make when reading selection 1
0
Did the reader attempt selection 4
Y
How many seconds did it take the reader to read selection 2
35
How many errors did the reader make when reading selection 2
1

Did the reader attempt selection 5
Y
How many seconds did it take the reader to read selection 3
42
How many errors did the reader make when reading selection 3
3
Did the reader attempt selection 6
Y
How many seconds did it take the reader to read selection 4
52
How many errors did the reader make when reading selection 4
4
Did the reader attempt selection 7
Y
How many seconds did it take the reader to read selection 5
98
How many errors did the reader make when reading selection 5
15
Did the reader attempt selection 8
Y
How many seconds did it take the reader to read selection 6
123
How many errors did the reader make when reading selection 6
17
Did the reader attempt selection 9
N
Did the student recognize most words instantly in
his grade level paragraph?
(answer NO if not attempted)
N
Did the student decode most words not known immediately when reading
at or above his grade placement?
(answer NO if not attempted)
N
Did the pattern of student miscues indicate
that visual discrimination was O.K.?
Y
Did the student appear to use context to adjust miscues?
N
Did the student use adequate phrasing during oral reading?
Y
Did the student read at an adequate rate of speed?
N
Did the student observe punctuation while reading?
Y
Did the student attempt to correct any miscalls made?
N
OConcerning the CLIPIR Silent Reading Measure:
Was the reading rate adequate in the grade level
paragraph? (answer NO if not attempted)
N
What was the number of the most difficult passage attempted by the
reader?
6

Concerning selection 2
What percentage of the main idea concepts did the reader give (enter the number only, e.g., 50% enter as 50)?
34
What % memories did the reader have in the sequential recall task (either in sequence or on further prompting)?
(again, enter the number only, e.g., 50% enter as 50)
14
What percentage of the inference question was answered correctly (again, only the number)?
0
What was the reading time in seconds?
12
Concerning selection 3
What percentage of the main idea concepts did the reader give (enter the number only, e.g., 50% enter as 50)?
100
What % memories did the reader have in the sequential recall task (either in sequence or on further prompting)?
(again, enter the number only, e.g., 50% enter as 50)
14
What percentage of the inference question was answered correctly (again, only the number)?
100
What was the reading time in seconds?
66
Concerning selection 4
What percentage of the main idea concepts did the reader give (enter the number only, e.g., 50% enter as 50)?
34
What % memories did the reader have in the sequential recall task (either in sequence or on further prompting)?
(again, enter the number only, e.g., 50% enter as 50)
9
What percentage of the inference question was answered correctly (again, only the number)?
100
What was the reading time in seconds?
100
Concerning selection 5
What percentage of the main idea concepts did the reader give (enter the number only, e.g., 50% enter as 50)?
34
What % memories did the reader have in the sequential recall task (either in sequence or on further prompting)?
(again, enter the number only, e.g., 50% enter as 50)
2
What percentage of the inference question was answered correctly (again, only the number)?
0
What was the reading time in seconds?
-99
Concerning selection 6
What percentage of the main idea concepts did the reader give (enter the

number only, e.g., 50% enter as 50)?
25
What % memories did the reader have in the sequential
recall task (either in sequence or on further prompting)?
(again, enter the number only, e.g., 50% enter as 50)
0
What percentage of the inference question was
answered correctly (again, only the number)?
0
What was the reading time in seconds?
71
OConcerning the CLIPIR Listening Comprehension Measure:
What was the number of the most difficult passage attempted by the
reader?
7
Concerning selection 2
What percentage of the main idea concepts did the reader give (enter the
number only, e.g., 50% enter as 50)?
34
What % memories did the reader have in the sequential
recall task (either in sequence or on further prompting)?
(again, enter the number only, e.g., 50% enter as 50)
21
What percentage of the inference quesion was
answered correctly (again, only the number)?
100
Concerning selection 3
What percentage of the main idea concepts did the reader give (enter the
number only, e.g., 50% enter as 50)?
100
What % memories did the reader have in the sequential
recall task (either in sequence or on further prompting)?
(again, enter the number only, e.g., 50% enter as 50)
29
What percentage of the inference quesion was
answered correctly (again, only the number)?
100
Concerning selection 4
What percentage of the main idea concepts did the reader give (enter the
number only, e.g., 50% enter as 50)?
25
What % memories did the reader have in the sequential
recall task (either in sequence or on further prompting)?
(again, enter the number only, e.g., 50% enter as 50)
50
What percentage of the inference quesion was
answered correctly (again, only the number)?
100
Concerning selection 5
What percentage of the main idea concepts did the reader give (enter the
number only, e.g., 50% enter as 50)?
34
What % memories did the reader have in the sequential
recall task (either in sequence or on further prompting)?

(again, enter the number only, e.g., 50% enter as 50)
30
What percentage of the inference quesion was
answered correctly (again, only the number)?
0
Concerning selection 6
What percentage of the main idea concepts did the reader give (enter the number only, e.g., 50% enter as 50)?
0
What % memories did the reader have in the sequential
recall task (either in sequence or on further prompting)?
(again, enter the number only, e.g., 50% enter as 50)
13
What percentage of the inference quesion was
answered correctly (again, only the number)?
0
Concerning selection 7
What percentage of the main idea concepts did the reader give (enter the number only, e.g., 50% enter as 50)?
50
What % memories did the reader have in the sequential
recall task (either in sequence or on further prompting)?
(again, enter the number only, e.g., 50% enter as 50)
22
What percentage of the inference quesion was
answered correctly (again, only the number)?
0
OConcerning the CLIPIR Vocabulary Measure:
how many words were correctly defined?
2
How many words were incorrectly defined?
5
How many word definitions were not attempted?
21

Concerning the CLIPIR Tester Debriefing Procedure:

Did the tester rate the reader's general sophistication of
language as "low"?
Y
Did the tester rate the student's overall affect as hostile, sullen or
flat?
N

Tester Debriefing Data Collection Complete.
DATE IS:JUL 8, 1981 TIME IS:15:57:43
INSTANT WORD RECOGNITION VITAL SIGN
-1.000
DECODED WORD RECOGNITION VITAL SIGN
-1.000
MEANING VOCABULARY VITAL SIGN
-1.000
ORAL READING VITAL SIGN
-1.000

SILENT READING COMPREHENSION VITAL SIGN
1.000
LANGUAGE EXPERIENCE VITAL SIGN
0.667
MOTIVATION TO READ AND LEARN TO READ VITAL SIGN
1.000
VISUAL DISCRIMINATION OF WORDS AND WORD SEGMENTS AND LETTERS
0.667
SOUND SYMBOL ASSOCATION FOR CONSONANT BLENDS
-1.000
SOUND SYMBOL ASSOCIATION FOR PHONOGRAMS
-1.000
ABILITY TO BLEND RECOGNIZED WORD PARTS
-1.000
ABILITY TO PARAPHRASE
-1.000
ABILITY TO DISTINGUISH MAIN IDEAS AND DETAILS
-1.000
ABILITY TO PREDICT MEANING FROM CONTEXT
-1.000
ABILITY TO DETERMINE CAUSALITY
0.0
PHRASING OF SPEECH DURING ORAL READING
1.000
READING RATE DURING ORAL READING
-1.000
READING RATE DURING SILENT READING
-1.000
VISUAL ACUITY
-1.000
RELEVANT ACADEMIC TIME ENGAGED ON READING TASKS
1.000
ADEQUACY OF SELF FEEDBACK DURING READING PRACTICE
-1.000
ATTITUDE TOWARD READING
1.000
INSIGHT PUNCTUATION GIVES SIGNALS FOR PROPER PHRASING
1.000
DATE IS:JUL 8, 1981 TIME IS:15:57:43
CLIPIR STANDARD INTERVIEW PROCEDURE
+1.000
IWR GRADE EQUIV CLIPIR INTERVIEW
+1.000
DWR GRADE EQUIV CLIPIR INTERVIEW
-1.000
ORAL READING CLIPIR INTERVIEW
-1.000
SILENT READING CLIPIR INTERVIEW
+1.000
ATTENTION MOTIVATION CLIPIR INTERVIEW
+1.000
USE OF PUNCTUATION CLIPIR INTERVIEW
-1.000
VISUAL ACUITY CLIPIR INTERVIEW

-1.000
AUDITORY ACUITY CLIPIR INTERVIEW
-1.000
TIME ENGAGED CLIPIR INTERVIEW
+1.000
CLIPIR SORT PROCEDURE
+1.000
IWR GRADE EQUIV CLIPIR SORT
-1.000
DWR GRADE EQUIV CLIPIR SORT
-1.000
WC GRADE EQUIV CLIPIR SORT
-1.000
USE OF INITIAL LETTERS FOR WORD RECOGNITION
+1.000
USE OF FINAL LETTERS FOR WORD RECOGNITION
+1.000
USE OF CONFIGURATION FOR WORD RECOGNITION
-1.000
VIS DISC CLIPIR SORT
+1.000
CLIPIR GATES MCKILLOP PROCEDURE
+1.000
SSA CONS BLEND CLIPIR GATES MCKILLOP
-1.000
SSA PHONOGRAMS CLIPIR GATES MCKILLOP
-1.000
BLENDING CLIPIR GATES MCKILLOP
-1.000
VIS DISC CLIPIR GATES MCKILLOP
+1.000
CLIPIR GRAY PROCEDURE
+1.000
OR GRADE EQUIV CLIPIR GRAY
-1.000
IWR GRADE EQUIV CLIPIR GRAY
-1.000
DWR GRADE EQUIV CLIPIR GRAY
-1.000
VIS DISC CLIPIR GRAY
-1.000
USE CONTEXT CLIPIR GRAY
-1.000
PHRASING CLIPIR GRAY
+1.000
ORAL READING RATE CLIPIR GRAY
-1.000
INSIGHT PUNCTUATION CLIPIR GRAY
+1.000
ORAL READING CORRECTION CLIPIR GRAY
-1.000
CLIPIR SILENT PROCEDURE
+1.000
SILENT READING RATE CLIPIR SILENT

-1.000
RC GRADE EQUIV CLIPIR SILENT
+1.000
MAIN IDEA CLIPIR SILENT
-1.000
CAUSALITY CLIPIR SILENT
-1.000
PARAPHRASE CLIPIR SILENT
-1.000
CLIPIR LISTENING PROCEDURE
+1.000
LC GRADE EQUIV CLIPIR LISTENING
+1.000
MAIN IDEA CLIPIR LISTENING
-1.000
CAUSALITY CLIPIR LISTENING
+1.000
PARAPHRASE CLIPIR LISTENING
-1.000
CLIPIR VOCABULARY PROCEDURE
+1.000
WC GRADE EQUIV CLIPIR VOCABULARY
-1.000
CLIPIR TESTER DEBRIEFING PROCEDURE
+1.000
LANGUAGE EXPERIENCE CLIPIR TESTER DEBRIEFING
-1.000
ATTITUDE CLIPIR TESTER DEBRIEFING
+1.000
INSTANT WORD RECOGNITION VITAL SIGN
-1.000
DECODED WORD RECOGNITION VITAL SIGN
-1.000
MEANING VOCABULARY VITAL SIGN
-1.000
ORAL READING VITAL SIGN
-1.000
SILENT READING COMPREHENSION VITAL SIGN
1.000
LANGUAGE EXPERIENCE VITAL SIGN
0.667
MOTIVATION TO READ AND LEARN TO READ VITAL SIGN
1.000
VISUAL DISCRIMINATION OF WORDS AND WORD SEGMENTS AND LETTERS
0.667
SOUND SYMBOL ASSOCATION FOR CONSONANT BLENDS
-1.000
SOUND SYMBOL ASSOCIATION FOR PHONOGRAMS
-1.000
ABILITY TO BLEND RECOGNIZED WORD PARTS
-1.000
ABILITY TO PARAPHRASE
-1.000
ABILITY TO DISTINGUISH MAIN IDEAS AND DETAILS

```
-1.000
ABILITY TO PREDICT MEANING FROM CONTEXT
-1.000
ABILITY TO DETERMINE CAUSALITY
0.0
PHRASING OF SPEECH DURING ORAL READING
1.000
READING RATE DURING ORAL READING
-1.000
READING RATE DURING SILENT READING
-1.000
VISUAL ACUITY
-1.000
RELEVANT ACADEMIC TIME ENGAGED ON READING TASKS
1.000
ADEQUACY OF SELF FEEDBACK DURING READING PRACTICE
-1.000
ATTITUDE TOWARD READING
1.000
INSIGHT PUNCTUATION GIVES SIGNALS FOR PROPER PHRASING
1.000
*$stop
```

Appendix D: Clinician Diagnostic Workup

A sample clinician diagnosis is included here. Such diagnoses were completed for a random sample of about one half of the pretested students (20). The diagnoses were based on the data from the student's protocol notebook without any identification of the student.

Does the student have a problem with INSTANT WORD RECOGNITION? (IWR)
(Circle One) Yes No
On what basis was this decision made?

1. *Describe one factor contributing to the problem with Instant Word Recognition.
   Deciding about the identity of a word based on one or two beginning letters.
   *Suggest remedial procedures for alleviating this factor.
   Approach all unknown or miscalled words from a DWR perspective (with some exceptions like "once"). See DWR 1.

2. *Describe another factor contributing to the problem with Instant Word Recognition.
   Inattention to all elements of the visual display, particularly medial phonograms.
   *Suggest remedial procedures for alleviating this factor.
   Highlight, by size, the salient portions of the word, e.g., blOUse. The ou digraph is a common one and can be transferred to other words subsequently.

3. *Describe another factor contributing to the problem with Instant Word Recognition.
   Overreliance on a store of relatively few highly redundant words common to reading materials in grades one through three.
   *Suggest remedial procedures for alleviating this factor.
   See 1 and 2 above.

4. *Describe another factor contributing to the problem with Instant Word Recognition.
   Insufficient extended reading practice which could act to expand store of instantly recognized words.
   *Suggest remedial procedures for alleviating this factor.
   See DWR #4.

5. *Describe another factor contributing to the problem with Instant Word Recognition.
   Auditory acuity problem.
   *Suggest remedial procedures for alleviating this factor.
   Screening with audiometer. Possible use of hearing aid or positioning student close to teacher so face to face contact can be maintained.

6. *Describe another factor contributing to the problem with Instant Word Recognition.
   Auditory discrimination problem.
   *Suggest remedial procedures for alleviating this factor.
   Wepman auditory discrimination test. If results positive, go to 5 above.

7. *Describe another factor contributing to the problem with Instant Word Recognition.

Visual acuity, fusion, tracking (etc.) problems.
*Suggest remedial procedures for alleviating this factor.
Vision screening with Keystone visual screening. If positive, refer to optometrist or opthalmologist for corrective treatment.

Does the student have a problem with DECODED WORD RECOGNITION? (DWR)
(Circle One) Yes No
On what basis was this decision made?

1. *Describe one factor contributing to the problem with Decoded Word Recognition.
Inadequate mastery of basic redundant sound-symbol associations.
*Suggest remedial procedures for alleviating this factor.
Use materials which begin with one-syllable words which incorporate the major redundant phonograms in isolation and in an extended contextual setting. The progression is from regular single vowel words (ship, black) to more irregular multiple voweled words (guided). Practice words in isolation and in extended contextual setting.

2. *Describe another factor contributing to the problem with Decoded Word Recognition.
Inability to reliably differentiate root words from bound morphemic affixes.
*Suggest remedial procedures for alleviating this factor.
Use materials which move from single syllable words to the affixed form such that practice is given in isolation and in an extended contextual setting. (star/starring; stare/staring; cut/cutter; cute/cuter; etc.)

3. *Describe another factor contributing to the problem with Decoded Word Recognition.
Absence of iterative strategies to (a) segment multisyllabic words, (b) transfer learned phonograms to each segment, (c) pronounce each segment, (d) recombine the segments into a spoken word, and (e) adjust the pronunciation if necessary.
*Suggest remedial procedures for alleviating this factor.
Use of materials which provide increaslingly lengthy multisyllabic words in isolation and in context, progressing from more to less regularity within segments.

4. *Describe another factor contributing to the problem with Decoded Word Recognition.
Insufficient extended reading practice with materials in a graduated sequence of difficulty (with respect to skills listed in factors 1, 2, and 3).
*Suggest remedial procedures for alleviating this factor.
Sufficient quantity of contextual material at every level in the hierarchy to consolidate recognition of major repetative phongrams.

5. *Describe another factor contributing to the problem with Decoded Word Recognition.
See IWR #5 (Auditory Acuity)
*Suggest remedial procedures for alleviating this factor.
See IWR #5 (Quditory Acuity)

6. *Describe another factor contributing to the problem with Decoded

Word Recognition.
See IWR #6 (Auditory Discrimination)
*Suggest remedial procedures for alleviating this factor.
See IWR #6 (Auditory Discrimination)

7. *Describe another factor contributing to the problem with Decoded Word Recognition.
See IWR #7 (Vision)
*Suggest remedial procedures for alleviating this factor.
See IWR #7 (Vision)

8. *Describe another factor contributing to the problem with Decoded Word Recognition.
Inadequate store of word meanings resulting in inability to adjust decoded word to correct spoken form.
*Suggest remedial procedures for alleviating this factor.
(1) Encourage use of context to support hypotheses about specific word meanings, (2) provide synonyms, actions, role playing, etc. to encourage educated guesses about word meaning.

Does the student have a problem with MEANING VOCABULARY? (MV)
(Circle One) Yes No
On what basis was this decision made?

1. *Describe one factor contributing to the problem with Meaning Vocabulary.
General sophistication of spoken language low, possibly due to inadequate or inappropriate language modelling.
*Suggest remedial procedures for alleviating this factor.
See DWR #8.

2. *Describe another factor contributing to the problem with Meaning Vocabulary.
Materials of sufficient complexity to enlarge semantic network are not accessible due to deficiencies in instant and decoded word recognition skills.
*Suggest remedial procedures for alleviating this factor.
See IWR/DWR.

3. *Describe another factor contributing to the problem with Meaning Vocabulary.
Ignorance of the meanings of common affixes, e.g., pre, sub, anti, inter, intra, un, de, dis, er, s, es, ed, etc.
*Suggest remedial procedures for alleviating this factor.
Always include meanings of affixed when doing decoding skills, and their effect on the root word.

4. *Describe another factor contributing to the problem with Meaning Vocabulary.
Ignorance of the etymology of common word parts, e.g., phono, bio, ology, graphy, micro, tele, etc.
*Suggest remedial procedures for alleviating this factor.
(None given by clinician)

5. *Describe another factor contributing to the problem with Meaning Vocabulary.
Insufficient practice in deducing word meaning from context.
*Suggest remedial procedures for alleviating this factor.
Provide a context for unknown words that is rich enough to support educated guesses about the meaning of the word in question.

Does the student have a problem with ORAL READING?(OR)
(Circle One) Yes No
On what basis was this decision made?

1. *Describe one factor contributing to the problem with Oral Reading.
   Instant word recognition inadequate.
   *Suggest remedial procedures for alleviating this factor.
   See IWR.

2. *Describe another factor contributing to the problem with Oral Reading.
   Decoded word recognition inadequate.
   *Suggest remedial procedures for alleviating this factor.
   See DWR.

3. *Describe another factor contributing to the problem with Oral Reading.
   Cognitive demands placed on reader to compensate for inadequate IWR/DWR skills results in insufficient attentional capacity to produce fluent, accurate translation from print to speech.
   *Suggest remedial procedures for alleviating this factor.
   Provide reading selections which parallel the sequence of skills built up from single to multisyllable and from more to less regular phonograms. Stress repeated reading for fluency and expressiveness. Use over-voicing with student. Model dramatic reading and have student repeat after you. Encourage self-evaluation: Does your reading sould less choppy? DOes it sound interesting to you? etc. Stress role of punctuation in regulating and enhancing fluent expression.

4. *Describe another factor contributing to the problem with Oral Reading.
   Inadequate performance in IWR/DWR results in slow, tedious and inaccurate reading which in turn discourages further practice.
   *Suggest remedial procedures for alleviating this factor.
   See #3 above.

5. *Describe another factor contributing to the problem with Oral Reading.
   Inability to use the sound system of the language to move beyond the single word stage.
   *Suggest remedial procedures for alleviating this factor.
   Repeated readings of short sections of text to a speed and accuracy criterion. This is a clinical procedure requiring one-on-one monitoring. Parental cooperation is usually necessary so student can practice at home.

6. *Describe another factor contributing to the problem with Oral Reading.
   Inappropriately rapid rate in the face of multiple miscues.
   *Suggest remedial procedures for alleviating this factor.
   See RX for DWR #1 and DWR #2. Stress accuracy over speed. Reward

stopping at unknown word for purposes of decoding.

Does the student have a problem with READING COMPREHENSION? (RC)
(Circle One) Yes No
On what basis was this decision made?

1. *Describe one factor contributing to the problem with Reading Comprehension.
   IWR inadequate.
   *Suggest remedial procedures for alleviating this factor.
   See IWR.

2. *Describe another factor contributing to the problem with Reading Comprehension.
   DWR inadequate.
   *Suggest remedial procedures for alleviating this factor.
   See DWR.

3. *Describe another factor contributing to the problem with Reading Comprehension.
   Limited store of word meanings available to student.
   *Suggest remedial procedures for alleviating this factor.
   See MV #3, MV #4, MV #5.

4. *Describe another factor contributing to the problem with Reading Comprehension.
   Dealing with difficulties in IWR/DWR/OR results in insufficient attentional capacity for systematic processing of incoming information.
   *Suggest remedial procedures for alleviating this factor.
   See IWR/DWR/OR remedial procedures.

5. *Describe another factor contributing to the problem with Reading Comprehension.
   Inadequate or inappropriate cognitive strategies for anticipating, chunking, and finally retrieving content read silently.
   *Suggest remedial procedures for alleviating this factor.
   Provide comprehension decision aids to be used for organizing incoming material and for retrieving it.

6. *Describe another factor contributing to the problem with Reading Comprehension.
   Inadequate volue of reading being done.
   *Suggest remedial procedures for alleviating this factor.
   Increase the amount of reading the student is required to read and organize/ both fiction and content area material.

7. *Describe another factor contributing to the problem with Reading Comprehension.
   Inability to reliably deal with inferential questions.
   *Suggest remedial procedures for alleviating this factor.
   Help student articulate what prior knowledge s.he already has about the material read and how to trust that knowledge to fill in the gaps made by inferential questions.

Does the student have a problem with LISTENING COMPREHENSION? (LC)
(Circle One) Yes No
On what basis was this decision made?

1. *Describe one factor contributing to the problem with Listening Comprehension.
   Limited store of word meanings.
   *Suggest remedial procedures for alleviating this factor.
   Increase the volume of reading the student is required to do.

2. *Describe another factor contributing to the problem with Listening Comprehension.
   Inadequate strategies for organizing and then retrieving the gist of what was heard.
   *Suggest remedial procedures for alleviating this factor.
   See RC #5. These strategies should be transferrable to the listening setting.

3. *Describe another factor contributing to the problem with Listening Comprehension.
   A gradual decrease in attention as the passage continues.
   *Suggest remedial procedures for alleviating this factor.
   This is more often a problem with the quality of the material being read than with the listener. Do not read material loaded with essentially unconnected detail, or materials that is poorly constructed in terms of logical structure and adequate redundancy.

4. *Describe another factor contributing to the problem with Listening Comprehension.
   A gradual decrease in attention as the session continues.
   *Suggest remedial procedures for alleviating this factor.
   Do not read all sections consecutively. When student begins to fidget, proceed to next task and pick up listening again later.

Does the student have a problem with ATTENTION / MOTIVATION? (A/M)
(Circle One) Yes No
On what basis was this decision made?

1. *Describe one factor contributing to the problem with Attention / Motivation.
   Repeated failure at performing reading tasks results in avoidance behaviors.
   *Suggest remedial procedures for alleviating this factor.
   See RX for DWR/OR/RC/WM. Provide continuing stream of informational feedback: why a performance was inadequate; why a performance was better; why it was good.

2. *Describe another factor contributing to the problem with Attention

/ Motivation.
Tediosness of performing reading tasks results in little or no practice of independent reading, contributing to further deterioration in performance and motivation.
*Suggest remedial procedures for alleviating this factor.
Provide time for sustained silent reading followed by oral reading and comprehension checks.

Appendix E: Equation of Clinician and Computer Diagnoses

The following pages contain lists of the following data in order:

1. A list of all diagnostic terms used by the clinician. The numbers preceeding the list correspond to the combined vocabulary statement under which this factor is included.

2. A list of the diagnostic terms used by the computer simulated clinician. The numbers preceeding the list correspond to the combined vocabulary statement under which this factor is included.

3. The combined vocabulary.

(1) Instant word recognition vital sign
(2) Decoded word recognition vital sign
(3) Meaning vocabulary vital sign
(4) Oral reading vital sign
(5) Silent reading vital sign
(6) Listening comprehension vital sign
(7) Attention/motivation vital sign
(8) Deciding about the identity of a word based on one or two beginning letters.
(10) Inattention to all elements of the visual display, particularly medial phonograms.
(43) Overreliance on a store of relatively few highly redundant words common to reading materials in grades one through three.
(1,38) Insufficient extended reading practice which could act to expand store of instantly recognized words.
(11) Auditory acuity problem.
(12) Auditory discrimination problem.
(13) Visual acuity, fusion, tracking (etc.) problems.
(17) Inadequate mastery of basic redundant sound-symbol associations.
(18) Inability to reliably differentiate root words from bound morphemic affixes.
(18,19,20) Absence of iterative strategies to (a) segment multisyllabic words, (b) transfer learned phonograms to each segment, (c) pronounce each segment, (d) recombine the segments into a spoken word, and (e) adjust the pronunciation if necessary.
(38) Insufficient extended reading practice with materials in a graduated sequence of difficulty (with respect to skills listed in factors 1, 2, and 3).
(21) General sophistication of spoken language low, possibly due to inadequate or inappropriate language modelling.
(1,2,3) Materials of sufficient complexity to enlarge semantic network are not accessible due to deficiencies in instant and decoded word recognition skills.
(22) Ignorance of the meanings of common affixes, e.g., pre, sub, anti, inter, intra, un, de, dis, er, s, es, ed, etc.
(23) Ignorance of the etymology of common word parts, e.g., phono, bio, ology, graphy, micro, tele, etc.
(37,38) Insufficient practice in deducing word meaning from context.
(1,2,25) Cognitive demands placed on reader to compensate for inadequate IWR/DWR skills results in insufficient attentional capacity to produce fluent, accurate translation from print to speech.
(1,2,24,38) Inadequate performance in IWR/DWR results in slow, tedious and inaccurate reading which in turn discourages further practice.
(17,18) Inability to use the sound system of the language to move beyond the single word stage.
(24) Inappropriately rapid rate in the face of multiple miscues.
(1,2,4,5) Dealing with difficulties in IWR/DWR/OR results in insufficient attentional capacity for systematic processing of incoming information.
(28,29,30) Inadequate or inappropriate cognitive strategies for anticipating, chunking, and finally retrieving content read silently.
(38) Inadequate volume of reading being done.
(30) Inability to reliably deal with inferential questions.
(31) Inadequate strategies for organizing and then retrieving the gist

of what was heard.
(40) A gradual decrease in attention as the passage continues.
(40) A gradual decrease in attention as the session continues.
(39,38) Repeated failure at performing reading tasks results in avoidance behaviors.
(42,38) Tediosness of performing reading tasks results in little or no practice of independent reading, contributing to further deterioration in performance and motivation.

(1) INSTANT WORD RECOGNITION VITAL SIGN
(2) DECODED WORD RECOGNITION VITAL SIGN
(3) MEANING VOCABULARY VITAL SIGN
(4) ORAL READING VITAL SIGN
(5) SILENT READING COMPREHENSION VITAL SIGN
(6) LANGUAGE EXPERIENCE VITAL SIGN
(7) MOTIVATION TO READ AND LEARN TO READ VITAL SIGN
(14) VISUAL DISCRIMINATION OF WORDS AND WORD SEGMENTS AND LETTERS
(15) SOUND SYMBOL ASSOCATION FOR CONSONANT BLENDS
(16) SOUND SYMBOL ASSOCIATION FOR PHONOGRAMS
(19) ABILITY TO BLEND RECOGNIZED WORD PARTS
(34) ABILITY TO PARAPHRASE
(35) ABILITY TO DISTINGUISH MAIN IDEAS AND DETAILS
(37) ABILITY TO PREDICT MEANING FROM CONTEXT
(36) ABILITY TO DETERMINE CAUSALITY
(25) PHRASING OF SPEECH DURING ORAL READING
(24) READING RATE DURING ORAL READING
(27) READING RATE DURING SILENT READING
(13) VISUAL ACUITY
(38) RELEVANT ACADEMIC TIME ENGAGED ON READING TASKS
(41) ADEQUACY OF SELF FEEDBACK DURING READING PRACTICE
(39) ATTITUDE TOWARD READING
(26) INSIGHT PUNCTUATION GIVES SIGNALS FOR PROPER PHRASING
(11) AUDITORY ACUITY CLIPIR INTERVIEW
(8) USE OF INITIAL LETTERS FOR WORD RECOGNITION
(9) USE OF FINAL LETTERS FOR WORD RECOGNITION
(10) USE OF CONFIGURATION FOR WORD RECOGNITION
(29) MAIN IDEA CLIPIR SILENT
(30) CAUSALITY CLIPIR SILENT
(28) PARAPHRASE CLIPIR SILENT
(32) MAIN IDEA CLIPIR LISTENING
(33) CAUSALITY CLIPIR LISTENING
(21) LANGUAGE EXPERIENCE CLIPIR TESTER DEBRIEFING
(39) ATTITUDE CLIPIR TESTER DEBRIEFING

(1) Instant word recognition
(2) Decoded word recognition
(3) Meaning Vocabulary
(4) Oral reading
(5) Silent reading
(6) Listening comprehension
(7) Attention/motivation
(8) Use of initial letters for word recognition
(9) Use of final letters for word reocgnition
(10) Use of configuration for word recognition
(11) Auditory acuity
(12) Auditory discrimination
(13) Visual acuity
(14) Visual discrimination
(15) Sound symbol association for consonant blends
(16) Sound symbol association for phonograms
(17) Sound symbol association general
(18) Segmentation
(19) Blending
(20) Adjustment
(21) Language experience
(22) Affix meanings
(23) Word part meanings
(24) Oral reading rate
(25) Oral reading fluency
(26) Use of punctuation in oral reading
(27) Silent reading rate
(28) Silent reading paraphrasing
(29) Silent reading main idea
(30) Silent reading inference
(31) Listening paraphrasing
(32) Listening main idea
(33) Listening inference
(34) General ability to paraphrase
(35) General ability to get main idea
(36) General ability to infer
(37) Ability to use context
(38) Relevant academic time engaged
(39) Attitude toward reading
(40) Attention span
(41) Self feedback
(42) Interest
(43) Other

Appendix F: Instruction Observation Form

Many iterations of the instruction observation form were used. Included here is the final version. The form consisted of two 8.5x14 inch pages taped together (reduced here, not proportionally). One form was used for each class. The most reliable information collected included type of task, time on task, materials, and instructor. Less reliable was the estimate of number of words or pages read in a task. Least reliable were measures of feedback on the second page, consisting of type and frequency of feedback from the instructor to the student. As can be seen, each entry in the form is divided into four parts one for each of four possible tasks in one class session. Rarely did more than four setting occur in one class session for one student.

Period 1

OBSERVER____

DATE____

	ENTIRE CLASS	CHAD	MICHAEL	BILLY	VIRGINIA	KEITH	RONDA
SETTING							
TIME							
MATERIALS							
INSTRUCTOR							
ASMT WORD READING							
ASMT ORAL READING							
ASMT SILENT READING							
ATTENTION							
	ARV.___ LEV.___ FLU.___	ARV.___ LEV.___ FLU.___	ARV.___ LEV.___ FLU.___	ARV.___ LEV.___ FLU.___	ARV.___ LEV.___ FLU.___	ARV.___ LEV.___ FLU.___	ARV.___ LEV.___ FLU.___

Period I

OBSERVER____

DATE____

	ENTIRE CLASS	CHAD	MICHAEL	BILLY	VONDIA	KEITH	RONDA
GSR VOWELS							
GSR PHONOGRAMS							
GSR CONS. CONS CLUST							
SEGMENTATION							
BLENDING							
ADJUSTMENT							
WORD RECOGNITION							
VOCABULARY							
PUNCTUATION							
CONTRACTIONS							
FLUENCY							
READING RATE							
INFLECTION							
COMPREHENSION							
USE OF SEMANTICS							
USE OF SYNTACTICS							
MAIN IDEA							
PARAPHRASE							
AFFECT FEEDBACK							
DISCIPLINE							
DISTRACTIONS							
NON-SPECIFIC COR. FB							
GLOBAL PERF. FEEDBACK							
HOMEWORK							
ATTENTION FEEDBACK							

Index

sim
sin
slo
sma
sma
sm
sn
so
sp
st
st
s
s
s
s
s
s
s
t
t
t
t

ta
ta

ta

ta

ta
ta
ta
ta
ta
ta
ta
t
t

t
t
t
t
t
t

t
t
t
t
t
t