5-2006

The Effect of Direct Instruction on the Generalization of Word Learning Skills for Students with Learning Disabilities

Brooke Ziwicki

This starred paper submitted by Brooke E. Ziwicki in partial fulfillment of the requirements for the Degree of Master of Science at St. Cloud State University is hereby approved by the final evaluation committee.

Chairperson

Dennis Nunts
Dean
School of Graduate Studies

THE EFFECT OF DIRECT INSTRUCTION ON THE GENERALIZATION OF WORD LEARNING SKILLS FOR STUDENTS WITH LEARNING DISABILITIES

by

Brooke Ziwicki

M.S., St. Cloud State University, 2006

A Starred Paper

Submitted to the Graduate Faculty

of

St. Cloud State University

in Partial Fulfillment of the Requirements

for the Degree

Master of Science

St. Cloud, Minnesota

May, 2006

TABLE OF CONTENTS

Chapter I

INTRODUCTION

Learning to read begins in early childhood and never stops. It is critical to the development of any other academic, social, behavioral, or life skill. Long-term consequences if a child does not learn to read can be devastating (Archer, Gleason, & Vachon, 2003). Unfortunately, the ability to read does not come as easily for some as it does for others. For as many as 20% of children, reading does not come automatically (Ellis, 1997). These are the students who are not in a position to discover for themselves the complex workings of the English language and that require instruction with a curriculum that emphasizes their individual differences and teaches them explicitly *what the rules of reading are* and *what they are not* (Tarver, 2000).

On a daily basis I teach *Reading Mastery* to a group of students with learning disabilities. This Direct Instruction program is designed to teach a multitude of reading skills, with earlier levels focusing on phonological structure and later levels focusing on more advanced reading skills (Engelmann & Brunner, 1988; Engelmann & Hanner, 1995). I see small improvements but wonder if the skills generalize to novel situations out of my classroom (i.e., academic reading, reading for enjoyment, etc.). Is there research to validate that I am teaching the right things in the correct

manner? Would a different approach prove more effective in helping students generalize reading skills from instructional to applied environments?

HISTORICAL BACKGROUND

The best way to teach reading has been a controversy for years, and at the heart of this controversy is the teaching of phonics (Drecktrah & Chiang, 1997). What has been considered effective instruction has changed a number of times during the last 30 years. In the 1960s and 1970s, effective instruction was based on large-scale studies such as First Grade Reading Studies (Bond & Dykstra, 1967) and Project Follow-Through Studies (Bock, Stebbins, & Proper, 1977; Stebbins, St. Pierre, Proper, Anderson, & Cerva, 1977).

During the early '80s, effectiveness was defined in terms of student outcomes. Along with the later part of the '80s and into the 1990s, came "best practices research" (Foorman & Torgesen, 2001). Researchers turned their attention to case studies of exemplary teachers and their technique of literature-based instruction. More widely known as whole language, this approach emphasizes independent reading and discussion of authentic literature. In programs underpinned by the whole language philosophy teachers approach skills through emphasizing meaningful context and place less emphasis on the direct and sequential instruction of phonics. The success of this technique is based on the assumption of reader response theory, which says that "readers play a central role in the construction of meaning and in social-constructivist theory" (Foorman & Torgesen, 2001, p. 204). More recently, this approach has been

combined with the traditional basal reader approach and is commonly referred to as the balanced literacy approach (Foorman & Torgesen, 2001).

Drecktrah and Chiang (1997) cited a longitudinal study (i.e., Juel, 1988) that supported the idea that without early teaching of decoding skills, students with poor reading skills in first grade where still poor readers in fourth grade. Another study indicated that 74% of students identified with a reading disability in third grade continued to show significant struggles with reading in ninth grade (Archer, Gleason, & Vachon, 2003). Therefore, researchers have suggested that at-risk readers often fail to discover the complex workings of the English language and need explicit and systematic phonics instruction (Archer, Gleason, & Vachon, 2003; Drecktrah & Chiang, 1997; Tarver, 2000). Educational researchers have also concluded that using a systematic approach with whatever is being taught is effective in improving educational opportunities for all students, especially, those with learning disabilities (Ellis, 1997).

Decoding and word recognition skills are the foundation of reading (Stanovich, 1996). "Poorly developed word recognition skills are believed to be the most pervasive and debilitating source of reading challenges" (Archer, Gleason, & Vachon, 2003, p. 90). If educators expect children to understand anything read on their own, youngsters must first correctly decipher print.

Readers must accomplish at least two things as their brains encounter written text. First, they must be able to recognize and read printed words, and second they need to construct meaning from text. If a child puts less energy into the first (i.e., a

child is more fluent), then they can apply more energy to understanding what they read (Archer, Gleason, & Vachon, 2003). No matter how old a person is, if the reader cannot decode fluently, she/he is not going to comprehend printed matter. Teachers' efforts need to be put into making certain that a child can decipher printed words. "The awareness of phonological structure of words is the best predictor of a child's reading success" (Drecktrah & Chiang, 1997, para. 4).

THEORETICAL BACKGROUND

Direct Instruction

In 1966, Siegfried Englemann, Carl Bereiter, and Wes Becker developed the theory of Direct Instruction while working with needy children (Kozioff, LaNunziata, Cowardin, & Bessellieu, 2001). The program originally carried the name DISTAR, an acronym for Direct Instruction System for Teaching Arithmetic and Reading, later changed to Direct Instruction System for Teaching and Remediation. In developing the program, Bereiter argued that a faster instruction pace was needed to help students with difficulties catch up with peers (Ryder, Burton, & Silberg, 2006). The Direct Instruction philosophy is that (a) teachers are responsible for learning, and (b) curriculum is the most important variable (Tarver, 2000). In-depth assessment of the program demonstrated that it is cognitively based and has been successful in teaching basic to advanced academic skills (Grossen, 2004).

As a model of cognitive learning, Direct Instruction has three components that are necessary for its effectiveness: analysis of knowledge systems, analysis of

communication, and analysis of behavior. Analysis of knowledge and communications is the *what* of the instruction and determines how the curriculum is designed by the publisher. The analysis of behavior is the *how* of the instruction and is what is delivered by the instructor (Tarver, 2000). The design and intent of direct instruction (in both its curricular and instructional aspects) is based on the assumption that if a student is not learning, the teacher isn't teaching. Of course, this assumes that educators apply the instruction and measurement system as it is outlined in the material. In other words, failure to learn results from inefficient and unsystematic instruction, and cannot be attributed to race, family background, socioeconomic status, or other factors (Kozioff, LaNunziata, Cowardin, & Bessellieu, 2001).

Direct Instruction, as developed by Carl Bereiter and Siegfried Engelmann (1966), consists of techniques and sequences that set standards according to what students can achieve. Commercial programs that contain these techniques are available for purchase by educators who are not directly trained by Engelmann and colleagues (Adams & Engelmann, 1996). Over 50 programs in reading, language, spelling, writing, mathematics, language, and science have been developed using their theory (Tarver, 2000). What Direct Instruction is not is teacher-directed instruction, or direct instruction. "Little 'd—i'", as it is often called, consist of behavioral descriptions that Rosenshine (1986) and Stallings (1987) developed based on Engelmann and colleagues' work at the University of Illinois (Adams & Engelmann, 1996).

In the 1970s and 1980s, research was conducted to determine the effective elements of Direct Instruction lessons (Slavin, 2000, p. 220). Field testing was done so that the average teacher could teach program skills with minimum training and still have all students be successful. DI programs are based on the following principles:

a. voluntary (i.e., operant) behavior is learned;
b. learning is dependent on the environment;
c. the teacher controls the environment and, thus controls voluntary behavior;
d. intelligent behavior is voluntary and therefore is learned and can be taught;
e. rate of learning is largely controlled by teaching technology (i.e., system);
f. successfully taught students have greater gains then students taught via other methods per unit of time;
g. thinking processes can first be taught as overt (usually verbal) processes;
h. the nature of the skill, not difference in the individual, is the logical determinant of a program's sequence;
i. when multiple interpretations might be learned, it is most efficient that the teacher that the teacher sequence the skills so that only one interpretation is learned;
j. it cannot be assumed that skills will transfer to related tasks unless the student is taught commonalities in the tasks;
k. quality of the instructional process is controlled by careful, systematic monitoring of student responses and feedback to the student; and
l. failure is a function of the instructional sequence, not the student. (Ryder, Burton, & Silberg, 2006, p. 180)

DI programs are also component driven:

a. small-group instruction;
b. oral responding in unison, which incorporates wait time and a respond signal to prevent higher performing students from answering before lower performing students;
c. rapid pacing with short breaks;
d. careful watching and listening to students' oral responses;
e. diagnosing and correcting errors through six steps (praise, model, lead, test, firm up, delayed test); and
m. motivation that begins as extrinsic reward—including physical contact, tickles, pats, handshakes—but which is gradually supplanted by intrinsic motivation. (Ryder, Burton, & Silberg, 2006, p. 180)

Tarver (2000) described Engelmann's Direct Instruction theory as a holistic approach in that it starts with the whole of knowledge, breaks it into logical components, and finally recombines it back to a whole knowledge system. Inherent in Direct Instruction is its ability to explicitly teach critical parts of a knowledge system as it relates to other parts. Tarver (2000) concluded that this whole-part-whole generalization technique represents the backbone of Direct Instruction. She further concluded that DI produces positive effects on student academic achievement. Perhaps Tarver's conclusion can be criticized on the basis that DI produces larger gains measured in traditional ways, perhaps not always attending to the higher-order comprehension skills that prove more difficult to measure.

Generalization

In dealing with generalization, a practitioner might wonder whether a student who reads "cat" in an instructional setting can also read the word in other important environments. One might pose a related mathematics-based question: When a student subtracts three apples from seven apples, can they also verbalize "4 apples" at the grocery store? These are questions related to the generalization of behaviors, or transfer of skills, from a teaching setting to other important environments. Generalization allows for students to apply the knowledge they have learned to a variety of situations. Children with learning disabilities are often unable to do this (Archer, Gleason, & Vachon, 2003; Tarver, 2000). Educational researchers in the field of learning disabilities, have cited four techniques, that when explicitly taught

through Direct Instruction, increase the chance of generalization (Archer, Gleason, & Vachon, 2003):

1. Students must be told when, why, and where they will use the strategy.
2. Students must become automatic in using the strategy.
3. Students must receive plenty of practice in using the strategy.
4. Students are explicitly told to use the strategy. (p. 95)

Engelmann developed DISTAR around three beliefs with the concept of generalization as the driving force. They are as follows: 1) generalization is taught initially rather than as a subset of a larger more concrete idea; 2) generalization needs to be taught explicitly and systematically, and; 3) generalization has been successfully demonstrated with skills taught via DI (Tarver, 2000).

Some researchers, however, view DI as a rote memory approach that, by its nature, will not produce transfer or generalization (Tarver, 2000). Adams and Engelmann (1996) devoted chapter three to refuting myths about DI, concluding that, "When Direct Instruction programs are used as intended, they facilitate responsive teaching" (p. 25). Given the controversial nature of initial reading instruction and the widespread use of Direct Instruction packages in special education, it is important to revalidate Adams's and Engelman's claims.

FOCUS OF THE REVIEW

Direct instruction programs for the delivery of code-based initial reading skills are widely used in teaching students with learning disabilities not only in the United States, but in Canada, the United Kingdom, and Australia (National Institute for Direct Instruction, 2005, p. 1). Although there are all sorts of programs available to teach the

content areas of reading, writing, and math, of most interest in this paper are reading programs that address word learning skills. Via a comprehensive review of literature, I first investigated whether or not word learning skills can be successfully taught via DI to students with learning disabilities. I then examine the effects, if any, of direct instruction on skill transfer. Finally, as part of the scope of this literature review, I determine to what degree that benefits of DI (if any) decrease, increase, or stay the same as students' age.

I reviewed articles that discussed the nature of Direct Instruction programs, their intended effects on generalization of skills, and their actual effect on generalization of skills. The use of comparative studies will be employed to help show differences between the effect that direct instruction programs have on skill transfer as compared to that of other methods or programs. With the exception of the extensive review *Direct Instruction: 25 years beyond DISTAR* and *Project Follow Through*, comparative studies that are not primarily related to word recognition and decoding are not reviewed. In addition to the peer-reviewed references chosen from scholarly journals, I reviewed teachers' guides from direct instruction programs to help describe the nature of programs and to seek documentation of transfer effects. The literature review covers materials published in the last decade.

DEFINITIONS

Direct Instruction (DI). A more extended definition of Direct Instruction is developed above. A succinct version reads as follows: "Direct instruction is a model

for teaching that emphasizes well-developed and carefully planned lessons designed around small learning increments and clearly defined and prescribed teaching tasks" (National Institute for Direct Instruction, 2005 p. 1). In this paper, I employ the term to mean Direct Instruction as defined above, but also DI of systematic phonics (defined below). DISTAR (Adams & Engelmann, 1996) programs and their more specific titles such as *Reading Mastery (*Engelmann & Hanner, 1995*)*, and all similar programs will be taken as fitting the definition of DI.

Generalization/Transfer. Generalization and transfer will be used interchangeably in this paper to discuss learning outcomes, although this usage is far from typical. Generalization, or the transfer of a behavior, is the degree to which an individual applies behaviors, skills, or concepts learned in one situation or set of situations to situations differing from the environmental conditions under which the skill was initially acquired. For example, if a student reads the words Japan, Italy, and San Francisco in their *Reading Mastery* textbook, do they read and understand them when encountered in the social studies text?

Maintenance. Maintenance is the continuation of a learned skill (Slavin, 2000, p. 157). For example, once a child learns to read, they show continued knowledge of that skill by reading and comprehending similar information presented to them over time. Some researchers have defined maintenance as generalization of a skill over the dimension of time (Maag, 1999).

Phonics/Systematic Phonics Instruction. In phonics instruction, educators specifically teach children about the "... relationships between letters (graphemes) and the individual sounds (phonemes) of spoken language" (Armbruster, Lehr, & Osborn, 2003, p. 12). For example, if a student understands the phonetic process, they will know that the sounds /b/, /a/, and /t/ form the word bat. "The goal of phonics instruction is to help children learn and use the alphabetic principle—the understanding that there are systematic and predictable relationships between written letters and spoken sounds" (Armbruster, Lehr, & Osborn, 2003, p. 12).

CONCLUSION

In Chapter I, it was established that many students in the U.S., especially those identified with learning disabilities (LD), experience difficulty learning to read. One widely-applied method with these students is Direct Instruction. The efficacy of Direct Instruction approaches for acquisition, maintenance, and generalization of reading skills by students with LD has not been reviewed in recent years. That is the thrust of Chapter II.

Chapter II

REVIEW OF LITERATURE

Effective instruction can not be judged solely on its ability to improve the performance of students, retention and transfer must also be considered (Adams & Engelmann, 1996). When as many as one out of five students struggle academically, educators must accelerate students' learning rates (Ellis, 1997). Unless educators prove able to teach more skills, at a faster rate, with more generalization in less time, struggling students will continue to perform at a lower standard than their peers (Adams & Engelmann, 1996). Direct Instruction (as defined in Chapter I), developed by Engelmann and colleagues in 1966, is a model purported to accelerate the reading acquisition in troubled readers.

In this chapter, I reviewed the effectiveness of DI in teaching word learning skills to students with learning disabilities. This is followed by a review of the relationship between Direct Instruction and reading-skill transfer. The final task was to find out the effect that transfer, if any, had on students at different age groups. Articles reviewed in Chapter II are arranged and selected using the following criteria: 1) articles reflecting the effectiveness of DI in producing skill acquisition, 2) articles reflecting transfer of DI skills at the elementary level, and 3) articles reflecting transfer of DI skills at the secondary level.

EFFECTIVENESS OF DIRECT INSTRUCTION IN PRODUCING INITIAL ACQUISITION OF READING SKILLS

In determining the effectiveness of Direct Instruction, Adams and Engelmann, (1996) completed a large-scale meta-analytic research review. Adams reviewed existing studies to select only the ones that fit the criteria. To ensure quality of the meta-analysis, a sample of 20 articles was reviewed by an additional research professor. The inter-rater reliability quotient was .94. Thirty-seven papers met initial criteria for the meta-analysis. From each research study the following information, Adams and Engelmann (1996) recorded the following information (representing independent variables):

Variable 1—Type of Student: These included those enrolled in general education, Follow Through, special education, or those who were evaluated during follow-up studies.

Variable 2—Year of Publication. Adams and Engelmann recorded the year that the research report was released.

Variable 3—Age/Grade of the Students: The age and/or grade of target students was recorded and divided into elementary, secondary and adulthood.

Variable 4—Subject: Adams and Engelmann recorded for analysis the subject matter addressed in the studies (e.g., mathematics, reading)

Variable 5—Type of Test: Assessment instruments employed to collect dependent variables (DV) were divided into the categories of norm-referenced (e.g., standardized achievement test) or criterion-referenced (e.g., teacher-made test).

Variable 6—Type of Research Design: Adams and Engelmann denoted whether investigators selected either experimental (with random assignment) or causal-comparative (quasi-experimental) designs.

Variable 7—Duration of Intervention: The authors recorded the length (in time) of the DI intervention.

> **Variable 8—Type of Teacher**: The types selected for comparison across studies were designated as either "experimental" (special teachers who presented the lessons just for the research project) or regular (the students' regular teachers).
>
> **Variable 9—Fidelity of Implementation**: The authors rated the papers in terms of whether or not researchers attempted to ensure that the intervention was conducted as described.
>
> **Variable 10—Where the study was conducted:** Adams and Engelmann recorded and analyzed whether the study was completed in the U.S. versus in some other country. (pp. 40-47)

Of the 37 studies that were originally identified, 34 involved active interventions of the DI model and fit the criteria with the 10 variables addressed above. From that group, 173 individual comparisons were made between a combination of regular education and Follow Through group, and the special education group. The general education studies and the Follow Through studies were combined because the post-intervention outcomes of the two were not statistically different. The studies were subjected to three different analyses: two polling procedures (polling of means and polling of statistically significant differences) and the meta-analysis (Adams & Engelmann, 1996).

The first set of analyses completed was a simple polling of means. The purpose of this analysis was to provide a measure of frequency and simply determined what percentage of the studies favored DI (in terms of measured outcomes), independently of how large such differences might be. Of the 174 comparisons, 151 (87.3%) favored DI while 21 (12.1%) favored non-DI instruction. One comparison tied (Adams & Engelmann, 1996).

A better indication of the magnitude of DI versus non-DI is calculated by a simple polling of statistically significant outcomes. Adams and Engelmann proposed that three possible outcomes could occur. One possibility was a statistically-significant result favoring DI. A second possibility was the existence of a statistically significant difference favoring non-DI. Third, there could have been no significant difference between the two groups. Three comparisons were excluded from this measure because they did not provide analysis of the DI model (Adams & Engelmann, 1996).

Two-thirds (64.1%) of the comparisons statistically significantly favored DI while only 1.2% of them favored non-DI instruction. "The Direct Instruction group achieved a statistically significant advantage 53.5 times more frequently than the non-Direct Instruction group" (Adams & Engelmann, 1996, p. 42). Although simplistic, the magnitude of this statistical difference is difficult to ignore.

The formal meta analysis. To determine effect size (effect of a particular intervention), the researchers divided mean differences by the standard deviation of the pooled population, thus producing a score roughly equivalent to a z score. Both effect sizes for the comparisons and effect sizes (expressed as a percentage of the standard deviation) for the individual studies were reported in the meta-analysis. The reason for this is that the 34 different investigation teams employed different mean scores, and possibly a different number of dependent variables. For example, one study used the *Metropolitan Achievement Test* (MAT) Total score while another used the *MAT* Word Attack scores, Comprehension scores, and the Total scores and yet

another team of researchers employed scores from the *Woodcock Johnson-Revised* Passage Comprehension and Reading Vocabulary tests. As a result, the different scoring procedures in each study might have resulted in a small difference in mean scores but a large difference in the number of effect sizes (Adams & Engelmann, 1996), thus the need for using the standard deviation of the pooled population.

Thirty-two of the 34 studies displayed positive effects for DI. The mean effect size per comparison was almost 1.0 while the mean effect size averaged per study was more than .75. This effect size is equivalent to a 3/4 of a standard deviation. In educational research, an intervention that changes a students behavior by 1/4 standard deviation is generally seen as both statistically- and clinically-significant. Therefore, educational interventions with effect sizes of .75 or more, 3/4 standard deviation, are rare events. Nineteen studies produced positive large effect sizes while 4 produced positive medium effect sizes and 4 produced positive small effect sizes. Of the 34 studies, 6 produced effect sizes that were non-significant while one study produced a negative small effect size (e.g., the non-DI subjects scored higher) (Adams & Engelmann, 1996).

Effect size for student category. Effect sizes were calculated for the type of student. The overall magnitude of this effect was large, .97 when calculated per comparison and .87 when calculated per study. The effect size per study for general education versus students with special needs were very similar (.82 for general education and .90 for special education) that the conclusion can be drawn that DI is an effective intervention for either type of student.

Effects as a function of time. The 32 studies chosen by Adams and Engelmann for meta-analytic treatment were published between 1972 and 1996. The outcomes was that the 6 studies from 1972-1980, 22 from 1981-1990, and 6 from 1991-1996 were consistent. The conclusion can be drawn that even though the DI curriculum underwent changes and revisions, it remained effective over that time period (Adams & Engelmann, 1996).

Grade level effects. Adams and Engelmann divided students (e.g., studies) into two grade-level groups (Elementary = K-5; Secondary/Adulthood = 6-adult). Though both groups improved, a t-test for group (on effect sizes) produced a statistically significant difference favoring the secondary group ($t = 3.33$, $df = 169$, $p<.001$). This may have been due to a relatively small sample size for the secondary group (33 comparisons for secondary and 140 comparisons for elementary).

Nine academic areas were subjected to effect size analysis per comparison. All nine resulted in effect sizes above the .25 that is needed to be considered educationally significant. Six academic areas had large effect sizes (social skills, math, spelling, health and science) while one had a medium effect size (reading), and two showed small effect sizes (legal concepts and language). To more readily understand this, the .69 effect size in DI reading was compared to a meta-analysis of whole language. Stahl and Miller's (1989) research study demonstrated a .09 effect size for whole language reading instruction. However, it should also be noted that a 1989 report by Slavin, Karweit, and Madden indicated an effect size of .07 for DI reading

comprehension and a .17 effect size for math problem solving. This information is not consistent with any other published reports (Adams & Englemann, 1996).

Type of independent variable. Adams and Engelmann compared norm-referenced and criterion-referenced measurements. Norm-referenced tests were those considered as standardized while criterion-referenced were those that were teacher made. Although both showed educationally significant effect sizes (ES for norm-referenced = .57 and criterion-referenced = 1.48), further analysis indicated that there was a statistically significant difference in results favoring criterion-referenced tests ($t = 6.59$, $df = 169$, $p < .001$).

Type of research design. The studies were divided into two categories of research (Adams & Engelmann, 1996). One type was causal-comparative, sometimes described as quasi-experimental (Gay, Mills, & Airasian, 2006). For this type, subjects are not randomly assigned and intact groups are used. The other type was experimental design. For this type, subjects are randomly assigned to intervention and control groups. In this meta-analysis, both designs produced large effect sizes. However, a t-test revealed a significant statistical difference between the two types of research with causal-comparative having a much larger effect (ES = 1.20) than did experimental (ES =.85) ($t = 2.30$, $df = 169$, $p < .05$).

The effect of learning to mastery. Adams and Engelmann (1996) explained that the concept of learning something to mastery is often listed as a "best practice." However, in 1987, Slavin analyzed mastery learning and found that when short term

studies were excluded from statistical analysis, mastery learning was minimal. Adams and Engelmann found that for DI, both effect sizes per comparison (up to 1 year = 1.10/over 1 year = .89) and effect sizes per study (up to 1 year = .95/over 1 year = .78) were large regardless if the intervention was less than one year or over a year.

Experimental teachers and general education teachers. The background and preparation of educators is important when deciding on a curriculum. If teachers must be highly trained, the usefulness of DI in general education settings may be limited. This meta-analysis showed that DI had a large effect size (.84-1.03) no matter what type of teacher taught the curriculum.

Fidelity of implementation. It is quite possible that an intervention in a study may not be implemented as designed. It is also quite common that studies implemented in a laboratory setting are monitored while those in the "real world" are not (Adams & Engelmann 1996). Twenty of the 34 research studies in this analysis had fidelity checks while 14 did not. The mean effect size for in vivo (naturalistic) studies was .74.

Location of intervention. Many of the research studies conducted on DI outside of the United States could not be included in Adams and Engelmann's meta-analysis because they did not meet entrance criteria. For example, some excellent studies done in Australia needed to be excluded because they did not contain a control group. Comparison effect size and effect size per study were both large in the United States,

but only the effect size per study was large in foreign investigations. The effect size per comparison for Non-United States was .67 which indicated a medium effect size.

Adams and Engelmann (1996 p. 48) concluded their DI meta-analysis with, "No matter which variable is analyzed, the results are overwhelmingly favorable (for DI)." These results represented many investigations over a wide geographic area, with many independent variables (e.g., sex, SES, special education status, country).

Project Follow Through

Over 10,000 students in 180 communities were chosen to participate in the Project Follow Through research study that originated from the Department of Education. Initially designed to be a program similar to Head Start, financial cutbacks changed the research study to a program evaluation from 1968 to 1976 instead of a service project. Part of the $500 million dollar project, was an evaluation of the effectiveness of nine different instructional approaches (one of which was DI) used with children in kindergarten through third grade. From 1976-1995, when funding was eliminated, it continued as a service project. "Project Follow Through was the largest, most expensive educational experiment ever conducted" (Adams & Engelmann, 1996, p. 67).

The first dimension of classifying the nine instructional approaches was sub-divided into three categories: the behavioral approach, the cognitive development approach, and the psychodynamic approach. Investigators divided the studies into three categories, "basic skills models, cognitive/conceptual skills models, and affective skills models" (Adams & Engelmann, 1996, p. 68). Second, ABT Associates

classified instructional approaches by their degree of structure. According to White (1973), a program can contain high structure, middle structure, or low structure. These were defined based on a teacher's ability to select materials according to their classroom needs. For example, a high structured program had predetermined roles for the teacher and student and a low structured program allowed teachers to choose curriculum based freely on their needs.

Basic skills models were those that were based on the belief that all behaviors are learned. The three programs included were: Behavior Analysis Model from the University of Kansas, Direct Instruction Model from the University of Oregon, and Language Development Model from Southwest Educational Development Laboratories. Behavior Analysis and Direct Instruction were classified as highly structured. The Language Development Model was classified as middle structure, a combination that involves low and high structure activities (Adams & Engelmann, 1996).

Cognitive/conceptual models were those that were based on normal cognitive growth. The three programs included were Cognitively-Orientated Curriculum Model from the High Scope Foundation, Florida Parent Education Model from the University of Florida, and Tucson Early Education Model (TEEM) from the University of Arizona. The TEEM and the cognitive curriculum were classified as medium structure while the structure of the parent education was unknown (Adams & Engelmann, 1996).

Affective skills models were those based on the socio-emotional development of a child. The three programs included were: Bank Street College Model from Bank Street College of Education, Open Education Model from the Education Development Center, and Responsive Education Model from the Far West Laboratory. The Bank Street curriculum and Open Education Model were classified as low structure and Responsive Education was rated as having medium structure (Adams & Engelmann, 1996).

Each model was evaluated across four to eight sites. The children needed to be starting kindergarten or first grade. The Follow Through school district then needed to identify a non-Follow-Through district to use as a comparison group. Just over 9,000 (9,255) third grade students participated in Follow Through schools while 6, 485 third grade students participated in the non-Follow-Through schools. Students' skills were tested at the beginning of the program in kindergarten and then in the spring of each year until third grade. Reliable and valid tests were selected in assessing the outcomes of each instructional approach. Data were collected in the areas of basic skills, cognitive and affective behaviors (Adams & Engelmann, 1996).

Data were analyzed by comparing the scores of each Follow Through group to the scores of each non-Follow-Through group. Researchers compiled non-Follow-Through scores into a pooled comparison group made up of all nine non- Follow Through models. If the effect size for the Follow Through model was about 1/4 standard deviation (.25 effect size) and was statistically significantly different from the non-Follow-Through group, a +1 was recorded. For the reverse situation (where non-

schools outperformed FT schools), a -1, was assigned if the non-Follow Through group had an effect size of .25. These +1 and -1 values were multiplied by 100, range of scores from -100 to +100 (Adams & Engelmann, 1996). A majority of the scores were negative. This meant that, on the average, the non-Follow-Through groups outperformed the Follow-Through groups. Comparisons were made on three types of skill outcomes: basic, cognitive, and affective. While the Open Education approach had the lowest scores in all three areas, Direct Instruction had the highest on all three classes of measurements.

An assumption might be made that the three approaches categorized as basic skills models would score highest on measurement of basic skills, approaches categorized as cognitive models highest on cognitive skills, and approaches categorized as affective models highest on affective skills. Results, however, diverged from this assumption. The authors ranked scores from 1 to 9 with 1 being the best rating and 9 being the worst. The basic skills models produced an average rank of 2.7 in basic skills, 2.8 in cognitive skills, and 2.7 in affective skills (Adams & Engelmann, 1996).

One would expect the cognitive models to achieve the best cognitive results. The average rank of the cognitive models (5.0) proved considerably worse than the basic skills rank of 2.8. Likewise, the affective models had the worst affective skills rank of 6.7 (compared with the basic skills rank of 2.7). These data does not support one frequent criticism of DI, specifically that it negatively affects students' affect.

The results of Project Follow Through underwent some reanalysis by other researchers. Bereiter and Kurland (1981) tightened criteria for inclusion in the study and then reanalyzed the different models. Overall, their reanalysis provided even stronger support for DI. Their research indicated that, "the Direct Instruction model results were vastly superior" (Adams & Engelmann, 1996, p. 78) to other forms of instruction. Becker and Carnine (1980) criticized the original model even more, tightening entrance criteria; however, the results showed that the highest scoring model, DI, scored even higher and the lowest scoring models, cognitive curriculum and parent education, scored even lower (Adams & Engelmann, 1996).

Educators want what is best for their students. At the same time they want to produce superior out-comes in all academic areas (Adams & Engelmann, 1996). In order to do so, the curriculum chosen must allow schools to "(a) implement the model across different school setting, (b) use the model across grade levels so that all students can be involved in a rigorous sequence of instruction, and (c) help students to feel good about themselves." The Follow Through data confirmed that Direct Instruction has these features. It further determined that "the amount of learning achieved per unit of time is probably twice as high for the Direct Instruction models as it is for non-Direct Instruction models" (Adams & Engelmann, 1996, p. 4).

Shippen, Houchins, Steventon, and Sartor (2005) compared two DI reading programs. They proposed to assess the effects of different reading programs on word reading efficiency and oral reading performance. Shippen et al. completed the study in an inner-city school in the Southeastern U.S.A. Seventh grade students who were 2 or

more years behind their grade level in reading were chosen for the study. Of the 55 students that participated, three were identified as LD while the other 52 were not disabled. The investigators did not assign a control group because the two groups were being compared to each other. Both programs were implemented at the same time with different groups of students. The DI treatment programs used were *Corrective Reading Decoding B2* (Engelmann, Johnson, Carnine, Meyer, Becker, & Eisele, 1999), *Corrective Reading Decoding C* (Engelmann, Meyer, Johnson, & Carnine, 1999), *and REWARDS* (Archer, Gleason, & Vachon 2000). Prior to the start of the intervention, four general education teachers were trained; two were trained in *Corrective Reading Decoding* while the other two were trained in REWARDS. Random sampling was used to ensure homogeneous grouping while teachers were also assessed to ensure procedural fidelity.

The investigators employed a quasi-experimental, pre-post test design. The *Test of Word Reading Efficiency* (Torgesen, Wagner, & Rashotte, 1997) and the *Gray Oral Reading Test* (Weiderholt & Bryant, 2001) were administered as pre- and post-test measures. The data was analyzed using a MANOVA to guarantee that individual differences were not a factor in the statistical analysis. Results indicated that regardless of program level (*Corrective Reading Decoding B2 or Corrective Reading Decoding C*) or program type (*Corrective Reading Decoding B2, Corrective Reading Decoding C, and REWARDS)* DI interventions significantly improved students' reading skills.

In addition to Program comparisons, a 2 x 4 within subject analysis was used to determine the effect of the independent variable time (e.g., pre vs. post) on the four dependent variables of word reading efficiency, reading rate, reading accuracy, and reading fluency. Regardless of the DI strategy implemented in the study, students who participated showed significant improvements in word reading efficiency, reading rate, reading accuracy, and reading fluency. This "confirms the effectiveness of highly structured, explicit, teacher-directed instruction for struggling readers" (Shippen, Houchins, Steventon, & Sartor, 2005, p. 180).

Drecktrah and Chiang (1997) surveyed teachers of general education students and teachers of students with learning disabilities within elementary settings in order to determine what instructional approaches (Direct Instruction, whole language, combination, etc.) educators actually used and the effects, if any, of programs. Secondary to this was to learn what influenced the teacher to choose the approach they did, and why they continued to use programs. Three-hundred surveys were mailed to licensed teacher in Wisconsin. The Department of Public Instruction randomly selected 100 second-grade teachers, 100 fifth-grade teachers, and 100 elementary teachers of students with learning disabilities. Drecktrah and Chiang (1997) argued that these teachers represented all Wisconsin elementary teachers.

Part 1 of the questionnaire dealt with teacher demographics. In Part II, teachers were asked to indicate the type of instruction they used. To help respondents perform the task, researchers provided a definition of whole language and direct instruction. A 5-point Likert-type scale was used to indicate the degree to which the

instruction was used in the classroom and to rate its effectiveness. Teachers were also asked to note why they used that particular approach and why they continued to use it. Part three contained 21 questions focusing on specific instructional strategies. Internal consistency was indicated by an Alpha co-efficient of .62 (Drecktrah & Chiang, 1997).

One hundred eighty-three surveys were returned with 59 of them from second-grade teachers, 70 of them from fifth-grade teachers, and 54 of them teachers of students with learning disabilities. Eighty-three percent of respondents were female, 64.9% had more than 10 years experience in teaching and over three-fourths (84.2%) were between 31 and 60 years of age (Drecktrah & Chiang, 1997).

Over 50% of respondents indicated the use of direct instruction in their classrooms. About three in five (61.1%) teachers of students with LD and 59.4% of second-grade teachers specified the use of Direct Instruction while only a slightly smaller percentage (58.6%) of fifth-grade teachers employed DI. However, 53.7% of teachers of students with learning disabilities reported that DI was effective and less than one-third of second- (30.5%) and fifth-grade (28.6%) teachers reported DI's effectiveness when used alone (Drecktrah & Chiang, 1997).

Of the teachers who reported the use of whole language (WL), the only group reporting over 50% (64.3%) use were fifth-grade teachers. However, less than half of those respondents rated WL as effective when used alone. Over one-third of teachers of students with learning disabilities indicated the use of whole language but only one-fourth of them indicated that it proved effective. Second grade teachers indicated the lowest percentage (12.7%) of use of WL. Drecktrah and Chiang (1997) found that

over 70% of all three respondent groups indicated that a combined approach of Direct Instruction and whole language is most effective when teaching reading and writing.

Kuder (1990) studied the performance of 48 students from seven self-contained, special education classroom. Each student was classified as "perceptually impaired," a term used to describe learning disabled students in New Jersey. The two groups were split evenly, 24 students in each group. The students were similar in age, sex, race, and Peabody Picture Vocabulary Test (Dunn & Dunn, 1981) results.

The teachers using DI were given appropriate training. They reported following the instructions of DISTAR in regard to grouping, time on task, and reinforcement. No other materials were used in the instruction of these children. A wide variety of basal reading series were used with the students in the traditional basal curriculum. Their instruction was supplemented with materials outside of their basal series (Kuder, 1990).

Reading achievement was measured with the Woodcock Reading Mastery Test (WRMT) (Woodcock, 1974). This norm-referenced test was given at the end of the first and second years of the DI program. Unfortunately, pre-intervention tests scores were not available; hence, the reason for the tests at the end of the second year. Comparisons were then made between the first and second year scores (Kuder, 1990).

At the end of the first year, no significant statistical differences were found between the DISTAR group and the non-DISTAR group. Investigators noted slight differences on the WRMT subtests in that the basal group performed slightly better on letter identification while the DISTAR group scored higher on the word attack and

word comprehension. Lack of pre-intervention WRMT scores made it difficult to support or not support the hypothesis of the effectiveness of DI with students with learning disabilities. It was quite possible that the LD students were significantly poorer readers to begin with and therefore, made greater progress. To further assess the hypothesis, comparison between scores at the end of the first year and second year were made (Kuder, 1990).

Not all 48 students from the original sample participated in the second year of the study. Eighteen students received their second year of DISTAR while only 8 of the original 24 received instruction via a basal reading series. Eight original "basal" students began receiving DISTAR instruction. Based on reading grade and reading age, both groups progressed. Although not statistically significantly so, the DISTAR group tended to make greater achievement gains than did non-DISTAR students (Kuder, 1990).

To assess changes in reading achievement, WRMT scores were also compared. Both DISTAR groups, the 2-year DISTAR group and the 1-year DISTAR/1-year basal group, made the most progress on the word identification and the word attack subtests. Kuder reported a statistically significant difference between those two groups and the basal-basal group on the word identification subtests ($F(2, 31) = 4.0$, $p < .05$) (Kuder, 1990). No other subtest scores indicated statistically significant differences (Kuder, 1990).

In general, the results of the research study did not support the author's hypothesis about DISTAR's effectiveness. Although, the researchers reported

improvement in word attack skills, the results indicated that the students continued to fall behind their peers in overall reading achievement. The results of this study must be viewed with caution as some design limitations were present. The most likely reason for the differences in outcomes was the population studied. Students included in this study were severely reading disabled. Students in previous studies were not as significantly disabled. This design limitation as well as the fact that the students came from low socioeconomic status and questionable teacher preparation, contribute to the validity of the results (Kuder, 1990).

The studies reviewed above suggest that Direct Instruction is effective in producing acquisition of initial and advanced reading skills in students with learning disabilities. However, once a student has acquired skills under one set of materials one might ask whether they experience success in others? Does a child's age at which they are introduced to DI affect their ability to transfer the knowledge? These questions are addressed below.

EFFECTIVENESS OF DI IN PRODUCING TRANSFER AT THE ELEMENTARY LEVEL

Lovett, Barron, and Benson (2003) explored generalization and transfer-of-learning by assigning children to one of two remedial reading programs or a control group. Both of the reading interventions were similar in targeting generalization of word learning skills but were different in how they addressed it.

One treatment group was instructed with the Phonological Analysis and Blending/Direct Instruction Program (PHAB/DI). This program consisted of lessons

from Englemann and colleague's development of DISTAR. The other treatment group underwent training with a program called Word Identification Strategy Training (WIST). This program, developed in their research laboratory, was based partially on the Benchmark School Word Identification/Vocabulary Development Program (Gaskins, Downer, & Gaskins, 1986).

Developers of both programs recognized a need for sub-syllabic segmentation during word identification learning and its significance in achieving transfer of learning to uninstructed words. However, they differed in the length of the sub-words they used to accomplish the goal and promote generalization. In the PHAB/DI program, researchers stressed the letter to sound relationship in the context of decoding while the WIST strategy focused on rime application and word identification strategies.

After 35 hours of training, both treatment groups' demonstrated significant improvements in their ability to generalize the word learning skills they were taught. Results of word reading probes varying in their "distance" from instructed words, and standardized and experimental test measures indicated significant improvements. However, while both programs were associated with large effect sizes, the different patterns of transfer that occurred indicated the existence of treatment-specific effects. Specifically, the PHAB/DI program showed transfer within the phonological skill domain, whereas the WIST program resulted in a broader-based generalization of real English words (i.e., generalization occurred on unknown words). Although, still not reading at grade level after remediation, students demonstrated marked improvements

and generalizations of letter-sound relationships, better decoding abilities and more accurate word recognition skills.

Lovett, Barron, and Benson (2003) compared strict phonological-based remediation to a combination of phonological and strategy training. Seventy hours of remediation instruction were given to 7-13 years olds in one of five program sequences: PHAB/DI and WIST; WIST and PHAB/DI; PHAB x 2; WIST x 2; or a control group of classroom survival skills and mathematics. Students were assessed five times during the intervention.

Student's who received the sequential treatment of PHAB/DI and WIST, demonstrated steeper learning curves than did the groups receiving treatment with only one program. Superior performance appeared in relation to non-word reading, letter-sound and keyword knowledge, and word identification measures of near and far transfer. Lovett, Barron, and Benson (2003, p. 283) suggested that while phonologically based intervention (i.e., DI) programs appear necessary to achieve gains, multiple treatments including strategy training produced more generalization. These findings demonstrated that transfer from nonsense-word decoding to other measures is best achieved with multiple, effective approaches (Lovett, Barron, & Benson, 2003, p. 281).

Ryder, Burton, and Silberg (2006) examined the effectiveness of DI compared to more traditional reading instructional approaches. The comparison included four approaches to reading instruction. The methods included cognitive apprenticeship

approaches, balanced reading, explicit explanation, and direct instruction (essentially DISTAR).

Investigators selected participants from the first through third grades in the Milwaukee Public Schools (MPS) and Franklin Public Schools (FPS). MPS students were considered urban while FPS students were considered suburban. With the exception of one FPS schools, all schools participated in the study for 3 years. MPS had two different treatment groups. In one school, educators employed DI solely while another school used a mixed-method approach (a combination of DI and other instructional methods). Teachers used the Houghton-Mifflin basal with the comparison group students. An administrative decision at FPS was made to implement DI as a compensatory program. Students were exposed to the general reading curriculum and then had an additional 30 minutes of DI everyday. These students were not identified as to whether they were eligible for special education services (Ryder, Burton, & Silberg, 2006).

To test for longitudinal differences between DI and non-DI students, the Gates-MacGinitie Reading Tests (MacGinitie, MacGinitie, Maria, & Drever, 2000) were administered at the start of the intervention and in the spring of each child's first, second, and third grade years. Statistical analysis was completed on composite reading and comprehension scores. Ryder, Burton, and Silberg (2006) conducted classroom observations and teacher interviews as part of the study. Results from these indicated that the characteristics of the educators themselves, more than the reading method alone, predicted reading achievement.

A two-way analysis of covariance (ANCOVA), with Method (of instruction) as the independent variable and posttest difference scores as the dependent variables, and pretest scores as the covariate; was calculated. Results indicated that DI and non-DI students differed significantly on mean difference (Post-test) scores. Non-DI students significantly outperformed their DI counterparts at every grade level. In grade 1 the difference in mean scores was 16.83 (SD = 6.54); the second grade difference in mean scores was 14.46 (SE 6.20); and the third grade difference in mean scores was 20.72 (SD = 6.95) (Ryder, Burton, & Silberg, 2006).

A two-way ANCOVA was also calculated to examine the consistency between school districts. Overall, this statistic indicated that FPS significantly outperformed MPS in reading achievement. With the exception of third grade students, non-DI students outperformed DI students in both school districts. The outcome of the interaction between third grade and method of instruction indicated that DI students at FPS outperformed all groups at that grade level. These results suggested that when teaching children to read, DI is an effective means with suburban students while non-DI methods appeared to be more effective with urban students (Ryder et al., 2006). This finding contradicted past studies that found DI to be a very effective approach in teaching reading to the economically disadvantaged child (Becker, 1977; Meyer, Gersten, & Gutkin, 1983).

Ryder et al. also calculated ANCOVA's comparing method of instruction (independent variable) on posttest comprehension scores (pretest scores served as a covariate). FPS-DI students outperformed all groups in terms of their comprehension

skills while MPS non-DI students outperformed all their groups in comprehension skills. However, regardless of the method students received or the district they attended, gains in comprehension decreased from grade 1 to grade 2. At the same time, the mean differences in scores indicated that the comprehension gains made from second to third grade years were not as significant for DI students as they were for non-DI students (Ryder, Burton, & Silberg, 2006).

Evidence from past research studies like Project Follow Through make DI an easy claim for educators that are seeking solutions for large numbers of students who need to improve their reading ability in a timely manner. Researchers suggested that although past scientific research is overwhelmingly in favor of DI, educators need to take into consideration its benefit for all students (Ryder, Burton, & Silberg, 2006).

When interpreting the findings of this study, limitations need to be considered. First, relatively small samples of students were available for the entire 3-year study. Changes in staffing, student assignment, and abandonment of the DI curriculum at one school after only 2 years are just a few of the problems that question this study's ability to generalize findings to classrooms and schools not a part of it. Second, the integrity of the DI programs was not accounted for. In other words, the authors did nothing to ensure the fidelity of the program. The effectiveness of DI in acquisition and generalization is dependent upon its execution (Adams & Engelmann, 1996). For example, if the direction is to re-read a passage if more than nine errors are made the first time, and the teacher does not comply, mastery is not assured, a centerpiece of direct instruction.

Compton, Olinghouse, Elleman, Vining, Appleton, Vail, and Summers (2005) implemented the PHAST strategy (Lovett, Lacerenza, & Borden, 2000) (lessons from the Direct Instruction approaches developed by Englemann and Brunner, 1988 and the Benchmark School Word Identification Program developed by Gaskins, Downer, and Gaskins 1986) to teach reading to 53 children with learning disabilities in grades three through five. In an attempt to accurately measure transfer of decoding skill acquisition, Compton et al. (2005) created a new approach to assessing individual responsiveness to instruction.

This responsiveness model was applied to assess individual differences in the transfer of decoding skills to other areas of reading acquisition. Because it is tied directly to the intervention, this assessment is more sensitive to change then other measures.

From the responsiveness measure an optimal growth curve was derived. This curve represented the number of words that could be decoded at different lesson intervals. A 50 word assessment that mirrored the optimal growth curve was developed using a multiple wave assessment to control for word difficulty. This was done so that performance on the responsive word list could be compared to the optimal growth curve. The researcher's objective was to develop a measure that might correctly assess individual differences in responses to instruction (Compton et al., 2005). This information was then used to test the hypothesis that children who experience the most decoding gains initially would be most likely to transfer their skills to other areas of reading (Bransford & Schwartz, 1999).

Students were randomly divided into triads. Because assessment of the PHAST intervention was the goal of the investigation, two groups for every one control group were assigned to PHAST strategy. The control groups received the standard reading curriculum provided by their special education teacher. In contrast, the PHAST groups received the first 60 lessons of the program taught by trained research assistants. To monitor responsiveness to the intervention, a 50 word list that mirrored the optimal growth curve for the PHAST program was administered at the beginning of the intervention, and then every twelve lessons. The Woodcock Reading Mastery Test-Revised (Woodcock, 1998), the Test of Word Reading Efficiency (Torgesen, Wagner, & Rashotte, 1997b), and the Gray Oral Reading Test-3 (Wiederholt & Bryant, 1992) were also used to assess intervention outcomes (Compton et.al., 2005).

The PHAST strategy was first evaluated to determine its effectiveness in improving the reading skills of students with learning disabilities. Differences in outcomes were statistically different on measures of word attack and word reading efficiency. Students in the treatment group improved from a mean score of 85.2 (SD 4.4) on the pre-test for word attack to a mean score of 90.3 (SD 6.4) on the posttest while students in the control group showed no sign of improvement in word attack with mean scores 87.4 (SD 7.6/8.2) both pre- and post-test. Similar results were observed with word efficiency (e.g., rate or fluency) scores (Compton et al., 2005).

The primary intent of this study was to explore the association between a student's individual response to the PHAST strategy and their ability to transfer

learned skills (i.e., the closer a child's individual growth curve was to the optimal learning curve, the greater the response to the intervention and, as a result, the greater the likelihood that the student would show signs of decoding skill acquisition in non-instructed words). The results indicated that the relationships between individual growth and optimal growth were significantly associated with gains in word attack, word identification, accuracy, fluency, and comprehension. The speed of these skills is what the experimental individual responsiveness procedure provided that would not have been available otherwise. "These results suggest that as gains on reading skills require a move from near to far transfer, the responsiveness parameters increase in their overall ability to predict variance in gain scores" (Compton et al., 2005, p. 64). The individual rather than the group is the appropriate unit of analysis when determining transferability of skills learned from a program. The researchers pointed to this as a possible reason for the inability of past studies to show transfer of decoding skill gains to other areas of reading.

Compton et al. (2005) provided support to Bransford and Schwartz's (1990) argument concerning transfer and original learning. That is, those students whose individual growth curves are most like the optimal growth curve were most likely to transfer decoding skill-gains to other reading tasks.

The experimental procedure that this research study is based on is promising. However, the researchers have some limitations to work though before it can properly be used to assess the learning of children. It is not a valuable piece of information to the common educational practitioner because of the relatively sophisticated statistical

skills required for implementation, interpretation, and analysis of the growth curves. To provide educators with a new procedure of evaluating children, the individual responsiveness measured must be simplified so it can be used efficiently.

EFFECTIVENESS OF DI IN PRODUCING TRANSFER STUDIES AT THE SECONDARY LEVEL

As students age, the material in school gets harder. The achievement gap, between those who struggled in elementary school and those who did not, begins to widen. Content area classes in secondary schools perpetuate the frustration of students who already feel inferior. Research studies show that providing intense, direct, and explicit instruction in reading is critical in closing the achievement gap that poor readers experience (Foorman, Francis, Beeler, Winkates, & Fletcher, 1997; Salinger, 2003). "By providing explicit, intense, and rule-based reading instruction, teachers increase the likelihood that older, poor readers will gain skills" (Shippen, Houchins, Steventon, & Sartor, 2005, p. 176). When used with LD students, DI provides inactive learners with intensive, sequenced, and teacher directed instruction.

Shippen et al. (2005) found that DI had more value for stronger readers. They stated that it is quite possible that DI provided progressive benefits for stronger readers while increasing the skills of naïve learners. This is consistent with findings reported by Compton et al. (2005), where the closer a child's individual growth curve was to the optimal learning curve, the greater the response to the intervention and as a result, the greater the likelihood the student would show signs of decoding-skill acquisition in other areas of reading.

In a social validity survey given to the participants in the seventh grade, 56% of the students agreed on a 3-point Likert-type scale, that DI helped them with reading in the other classes. Sixty-seven percent agreed that DI helped them to become a better reader while 70% thought DI would be helpful to other middle school students.

Grossen (2004) carried out a comprehensive study of DI programs at a secondary level with high-risk students who were performing at least two levels below the grade in which they were enrolled. She chose students who scored at the 1st percentile on The Multi-Level Academic Survey Test (MAST) reading pre-test to evaluate the effects of the DI model with low-achieving students. Post-test results illustrated significant gains. Of the 39 students who scored at the 1st percentile on the pre-test, only 11 remained there at post-test time. These results were similar to students who were not identified as low achievers. The (MAST) reading results from year one was compared to results from the same reading test the second year. The median percentile moved from 9 to 25 (Grossen, 2004).

SUMMARY

In Chapter II, articles were reviewed that reflected the effectiveness of DI in producing skill acquisition and the relationship between direct instruction and reading skill transfer. It was established that DI is effective in teaching students to read using systematic and sequenced lessons. Additional evidence was cited that supported transfer of the reading skills when the fidelity of the programs is held in tact.

Chapter III

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

There is little debate about the importance of reading. However, while some children come to school having already acquired the prerequisite skills necessary to acquire formal reading skills, many do not. As many as 10 million children in the United States are affected by reading problems (Shippen, Houchins, Steventon, & Sartor, 2005). Researchers have spent many hours, years, and money searching for the answer to this question: What is the most effective and efficient way to teach students to read? The answer is inconclusive. Until the recent development of balanced literacy, the instructional approaches that have dominated tended to be either skill- or meaning-based (Pressley, 1998).

The focus of this literature review was the effectiveness of Direct Instruction (i.e., a skills-based approach) for students with learning disabilities, or related reading difficulties. The weight of the evidence supported DI's effectiveness, thus raising questions about maintenance and generalization (transfer) of acquired skills. Also, it was important to determine whether the effects of DI differed by age. Unfortunately, insufficient data existed to draw a conclusion on the latter issue.

DI effectiveness. Adams and Engelmann's (1996) meta analyses provided overwhelming evidence in support of DI's effectiveness. For each of the 10 variables analyzed, impressive effect sizes running from .75 to 1.0 occurred. In some cases, effect sizes over 1.0 were observed. Keeping in mind, that in educational research an effect size of .25 is considered statistically significant, the results provided insight to the enormous impact DI can have on the acquisition of basic reading skills and comprehension in troubled readers. Likewise, in the *Project Follow-Through* study (Adams & Engelmann, 1996), schools employing basic skills models (including generic direct instruction and DISTAR), consistently out-performed their counterparts employing *any* other method. DI students achieved at higher levels than did comparison students on basic, cognitive, and affective skills. More modern, balanced programs, it must be acknowledged, were not widely disseminated at the point where Adams and Engelmann were comparing other methods to DI.

One variable from the DI meta-analysis that provided direct insight into the likelihood of the transfer of skills out of the program was the type of measurements used to test skill acquisition. Criterion-referenced tests showed a large effect size of 1.48 while norm-referenced tests showed a medium effect size of .57. Although both educationally significant, further statistical analysis revealed results favoring criterion-referenced tests. This is important because students were not as likely to transfer their skills (or possess the right skills) when more indirect measures of achievements were employed as they were when they participated in teacher-made tests (based more directly on content begin taught) (Adams & Engelmann, 1996).

Kuder's reached results that differed from those reported in the meta analysis by Adams and Engelmann. In specifically dealing with students with learning disabilities, the results indicated that these students continued to fall behind their peers academically. However, limitations caused the study to be interpreted cautiously. Problems included low socio-economic status, questionable teacher preparation, and severity of disability. This was the case in other research studies that did not have the same outcomes as Adams and Engelmann's meta-analysis.

Effectiveness of DI in producing transfer at the elementary level. Although a vast majority of researchers have found DI effective, many studies have failed to find or report transfer. Compton, Olinghouse, Elleman, Vining, Appleton, Vail, and Summers (2005) found that elementary-aged students, who fall along the optimal growth curve for a program, are the ones that have the most success with the program. They created an experimental analysis to determine this and concluded with the need for the procedure to be simplified if it is to be useful to the common practitioner. Their results seem to support the practice of stressing reading efficiency (rate/ fluency) during early learning phases.

Effectiveness of DI in producing transfer at the secondary level. Shippen, Houchins, Steventon, and Sartor (2005) studied the reading performance of seventh graders served in a middle school setting. They agreed with the elementary findings from Compton et al. (2005) that DI exerted stronger effects on students with superior reading skills. This evidence also supported Kuder's (1990) claim that DI was not as

effective with students with learning disabilities as it is for students with no special needs (as is true with virtually any reading approach). However, Shippen, et al. found that 56% of the students in their investigation suggested improvement in other classes as a result of reading instruction in the resource room.

Generalization and transfer. Though the evidence for the effectiveness of DI in producing skill acquisition was overwhelming, less clear evidence exists for skill transfer—especially to grade-level materials. Many teachers of students with learning disabilities try to integrate students into the standard curriculum. Studies have failed to conclusively validate DI's effectiveness in producing transfer to grade-level material. However, Compton et al. (2005) provided some evidence supporting of transfer of learned DI skills. At the same time, the program used as the intervention for the treatment group in that study, the PHAST strategy, was one created out of the research conducted by the team at The Hospital for Sick Children (Lovett, Barron, & Benson, 2003).

This strategy begins with phonological and letter-sound training (DI) and then uses it as a framework for teaching word identification strategy methods. This is quite different from straight DI approaches such as *Reading Mastery*. The findings from these two studies (Compton et al., 2005; Lovett, Barron, & Benson, 2003) suggest that the "optimal approach in instructing children with learning disabilities was a combined intervention model that included both direct instruction and strategy instruction methods" (Lovett, Barron, & Benson, 2003, p. 283). Although this is valid research,

these results are not conclusive enough to overturn years of overwhelmingly positive support of DI research.

Some children will always come to school prepared and ready to become effective and efficient readers. However, there will also always be children who are not. When this 20% of students walk through the classroom door, educators, have a responsibility to teach them to the best of their ability (Ellis, 1997). Ryder, Burton, and Silberg (2006 p. 190) argued that, "research shows that no single approach works for all children." For students who are not in a position to discover for themselves the complex workings of the English language, Direct Instruction will always be a viable resource. It will teach them the "*what is*" and "*what is not*" so that they too, can be as successful in school as their academic counterparts (Tarver, 2000). In the end, teachers will need to match instruction and materials with the characteristics of their students through careful assessment. No "magic bullet" probably exists for all children—even DI.

RECOMMENDATIONS

A reading of the existing research on direct instruction generates recommendations in two domains. First, recommendations for current best practices are offered. This is followed by recommendations for future research, primarily a call for better demonstration of near and far generalization.

Recommendations for practice. Over 50 studies (if all 34 meta-analysis studies are included) supported: (1) DI for the acquisition of reading skills, and (2) the rate of learning as the best predictor of transfer to other words, especially in fluency and word attack skills.

At this time, no other method of teaching reading surpassed DI in transfer, so until better data exists, the following recommendations are warranted:

1. Teacher's of students with reading difficulties should teach skills in a structured manner, such as recommended in DI.
2. Teacher's of students with reading difficulties should emphasize phonics and "breaking the code."
3. Educators working with students with reading difficulties should emphasize fluency.
4. Teachers of students with reading difficulties should plan for transfer in the manner outlined in Chapter I of this paper.

An added recommendation is that educators ought to consider multiple approaches based on students' individual needs, as a few studies have brought a "DI-only" approach into some question.

RECOMMENDATIONS FOR FUTURE RESEARCH

Research surrounding the effectiveness of DI was completed at length in the seventies and eighties. Most of that research showed DI with positive effects. However, children are changing. Their abilities to be academically successful are at

greater risk of academic failure than they were when the construct of DI was developed. For example, there are more two-parent working families and single parent homes (National Research Council and Institute of Medicine, 2000) than there were in the seventies, causing less time for children to access their parents for help at home . Children are being diagnosed with more significant and new disabilities that need to be addressed (Ellis, 1997). As a result, it is time for new research to be conducted to establish the effectiveness of Direct Instruction approaches with students of the 21st century.

As Kuder pointed out in 1990, research does not firmly establish the effect DI has on students with learning disabilities. No conclusive data existed since that point either. In addition to that, studies need to be completed that address the issue of generalization and transfer as new DI programs become available (e.g., REWARDS). I would recommend, that because of the number of teachers, both special education and general education, using DI with this type of student, research monies be put into addressing this issue further.

REFERENCES

REFERENCES

Adams, G. L., & Engelmann, S. (1996). *Research on direct instruction: 25 years beyond DISTAR.* Seattle WA: Educational Achievement Systems.

Archer, A. L., Gleason, M. M., & Vachon, V. L. (2000). *REWARDS: Reading excellance: Word attack and rate development strategies.* Longmount, CO: Sopris West.

Archer, A. L., Gleason, M. M., & Vachon, V. L. (2003). Decoding and fluency: Foundation skills for struggling older readers. *Learning Disability Quarterly, 26*(2), 89-101.

Armbruster, B. B., Lehr, F., & Osborn, J. (2003). *Put reading first: The research building blocks for teaching children to read.* Jessup, MD: ED Pubs.

Becker, W. C. (1977). Teaching reading and language to the disadvantaged: What we have learned from field research. *Harvard Educational Review, 47,* 518-543.

Becker, W. C., & Carnine, D. W. (1980). Direct instruction: An effective approach to educational intervention with disadvantaged and low performers. In B. B. Lahey & A. E. Kazdin (Eds.), *Advances in clinical child psychology* (pp. 429-473). New York: Plenum.

Bereiter, C., & Engelmann, S. (1966). *Teaching disadvantaged children in the preschool.* Engelwood Cliffs, N J: Prentice-Hall.

Bereiter, C., & Kurland, M. (1981-1982). A constructive look at Follow Through results. *Interchange, 12,* 1-22.

Bock, G., Stebbins, L. B., & Proper, L. (1977). *Education as experimentation: A planned variation model. Vol. IV-B, effects of follow through models.* Cambridge, MA: ABT Associates.

Bond, G., & Dykstra, R. (1967). The cooperative research program in first-grade-reading instruction. *Reading Research Quarterly, 2,* 1-142.

Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. In Iran-Nejad & P. D. Pearson (Eds.), *Review of research in education* (pp. 61-100). Washington, D.C.: American Educational Research Association.

Compton, D. L., Olinghouse, J. G., Elleman, A., Vining, J., Appleton, A. C., Vail, J., & Summers, M. (2005). Putting transfer back on trial: Modeling individual differences in the transfer of decoding-skill gains to other aspects of reading acquisition. *Journal of Educational Psychology, 97*(1), 55-69.

Drecktrah, M. E., & Chiang, B. (1997). Instructional strategies used by general educators and teachers of students with learning disabilities. *Remedial and Special Education, 18*(3), 174-181.

Dunn, L., & Dunn, L., (1981). *Peabody picture vocabulary test-Revised.* Circle Pines, MN: American Guidance Service.

Ellis, W. (1997). Interventions for students with learning disabilities. *NICHCY News Digest, 25,* 2-17.

Engelmann, S., & Bruner, E. C. (1988). *Reading mastery I/II fast cycle: Teacher's guide.* Chicago, IL: Science Research Associations.

Engelmann, S., & Hanner, S., (1995). *Reading mastery III: Teacher's guide.* Columbus, OH: SRA/McGraw-Hill.

Engelmann, S., Johnson, G., Carnine, L., Meyer, L., Becker, W., & Eisele, J. (1999). *Corrective reading decoding strategies B2.* Columbus, OH: SRA.

Engelmann, S., Meyer, L., Johnson, G., & Carnine, L. (1999). *Corrective reading decoding skill applications C.* Columbus, OH: SRA.

Foorman, B., Francis, D., Beeler, T., Winkates, D., & Fletcher, J. (1997). Early intervention for children with reading problems: Study designs and preliminary findings. *Learning Disabilities, 8*(1), 63-71.

Foorman, B. R., & Torgesen, J. (2001) Critical elements of classroom and small-group instruction promote reading success in all children. *Learning Disabilities Research & Practice, 16*(4), 203-212.

Gaskins, I. W., Downer, M. A., & Gaskins, R. W. (1986). *Introduction to the benchmark school word identification/vocabulary development program.* Media, PA: Benchmark School.

Grossen, B. (2004). Success of a direct instruction model at a secondary level school with high-risk students. *Reading & Writing Quarterly, 20,* 161-178.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grade. *Journal of Educational Psychology, 80,* 437-447.

Kozioff, M. A., LaNunziata, L., Cowardin, J., & Bessellieu, F. B. (2001). Direct Instruction: It's contributions to high school achievement. *High School Journal, 84,* 54-72.

Kuder, S. J. (1990). Effectiveness of the DISTAR reading program for children with learning disabilities. *Journal of Learning Disabilities, 23*(1), 69-71.

Lovett, M. W., Barron, R. W., & Benson, N. J. (2003). Effective remediation of word identification and decoding difficulties in school-age children with reading disabilities. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of Learning Disabilities* (pp. 273-292). New York, NY: The Guilford Press.

Lovett, M. W., Lacerenza, L., & Borden, S. (2000). Putting struggling readers on the PHAST track: A program to integrate phonological and strategy-based remedial reading instruction and maximize outcomes. *Journal of Learning Disabilities, 33,* 458-476.

MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Drever, L. G. (2000). *Gates-MacGinitie reading tests.* Itaska, IL: Riverside.

Magg, J. W. (1999). *Behavior management: From theoretical implications to practical applications.* Belmont, CA: Wadsworth Group/Thomson Learning.

Meyer, L. A., Gersten, R. M., & Gutkin, J. (1983). Direct Instruction: A project follow through success story in an inner-city school. *Elementary School Journal, 84,* 241-252.

National Institute for Direct Instruction. (2005). *What is direct instruction (DI)?* Retrieved November 14, 2005, from http://www.nfidi.org.

National Research Council and Institute of Medicine. (2000). After-school programs to promote child and adolescent development: Summary of a workshop: Committee on community-level programs for youth. In J. A. Gootman, (Ed.), *Board on children, youth, and families, commission on behavioral and social sciences and education.* Washington, DC: National Academy Press.

Pressley, M. (1998). *Reading instruction that works: The case for balanced teaching.* New York, NY: The Guillford Press.

Rosenshine, B. (1986). Synthesis of research on explicit teaching. *Educational Leadership, 43,* 60-69.

Ryder, R. J., Burton, J. L., & Silberg, A. (2006). Longitudinal study of direct instruction effects from first through third grade. *The Journal of Educational Research, 99*(3), 179-191.

Salinger, T. (2003). Helping older, struggling readers. *Preventing School Failure, 47*(2), 79-85.

Schloss, P. J., & Smith, M. A. (1998). *Applied behavior analysis in the classroom* (2nd ed.). Boston, MA: Allyn and Bacon.

Shippen, M. E., Houchins, D. E., Steventon, C., & Sartor, D. (2005). A comparison of two Direct Instruction reading programs for urban middle school students. *Remedial and Special Education, 26*(3), 175-182.

Slavin, R. E. (1987). Mastery learning reconsidered. *Review of Educational Research, 57,* 175-213.

Slavin, R. E. (2000). *Educational psychologist: Theory and practice* (6th ed.). Needham Heights, MA: Allyn and Bacon.

Slavin, R. E., Karweit, N. L., & Madden, N. A. (1989). *Effective programs for students at risk.* Boston: Allyn and Bacon.

Stahl, S. A., & Miller, P. D. (1989). Whole language and language experience approaches for beginning reading: A quantitative research synthesis. *Review of Educational Research, 59,* 87-116.

Stallings, J. (1987). Longitudinal findings for early childhood programs: Focus on Direct Instruction (p. 10). (ERIC Document Reproduction Services No. 297 874)

Stanovick, K. E. (1996) Word recognition: Changing perspectives. In R. Barr, M. L. Kamil, P. B. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (pp. 418-452). New York: Longman.

Stebbins, L., St. Pierre, R., Proper, E., Anderson, R., & Cerva, T. (1977). *Education as experimentation: A planned variation model, an evaluation of follow through, Vol. IV-A.* Cambridge, MA: ABT Associates.

Tarver, S. (2000). Direct Instruction: Teaching for generalization, application and integration of knowledge. *Learning Disabilities, 10*(4), 201-207.

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1997b). *Test of word reading efficiency.* Austin, TX: Pro-Ed.

White, S. H. (1973). *Federal programs for young children: Review and recommendations.* (HEW Service Publication No. (OS) 74-101). Washington, DC: US Government Printing.

Wiederholt, L., & Bryant, B. (1992). *Gray oral reading test—3.* Austin, TX: Pro-Ed.

Woodcock, R. W. (1974). *Manual for the Woodcock reading mastery test.* Circle Pines, MN: American Guidance Service.

Woodcock, R. W. (1998). *Woodcock reading mastery test—Revised/normative update.* Circle Pines, MN: AGS.