A META-ANALYSIS COMPARING ALTERNATIVE METHODS OF INDIVIDUALIZED AND TRADITIONAL INSTRUCTION IN SCIENCE

by

Nancy Carol Aiello

Dissertation submitted to the Graduate Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Educational Research and Evaluation

APPROVED:

Lee M. Wolfle, Chairman

Lawrence H. Cross

Maxine A. Enderlein

Marilyn Lichtman

J. Clarence Sasscer

January, 1981

Blacksburg, Virginia

DEDICATION

This manuscript is dedicated to my parents, and the late . Their trust and belief in me allowed me to grow.

ACKNOWLEDGMENTS

There are many people who have assisted me in conducting this study. Dr. Lee M. Wolfle, my advisor, was not only my main source of professional guidance, but also provided understanding and support. He gave willingly of his time and energies and was a very gentle and patient advisor. The other members of my committee, Drs. Cross, Enderlein, Lichtman and Sasscer, were each helpful and supportive. The guidance they gave is appreciated.

This program of study could not have been completed without the help of , Provost of the Loudoun Campus of Northern Virginia Community College. He provided encouragement and support in many ways, including two summers of academic leave for full time study.

, Instructional Assistant at the Loudoun Campus, provided coaching on the use of the word processor. Her help is very much appreciated.

A very special thanks goes to my son , who waited for what must have seemed an eternity for the completion of this dissertation. Thank you, , for being so patient.

The personal encouragement and support given by made the completion of this doctorate a reality. Thank you.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

INTRODUCTION

This study statistically integrates research studies conducted on alternative methods of individualized instruction. Individualized instruction is designed to meet the unique abilities, goals, and learning rates of individual students (Triezenberg and McLeod, 1972, p.6). It has been used in educational communities for nearly a century. The first recorded account of individualized instruction was in 1888 in Pueblo, Colorodo (Billett, 1932, p. 289), and probably dealt with an extremely uniform student group. In the early 20th century, compulsory education, coupled with increased urbanization and immigration, resulted in a more diverse student population. This diversity spurred an awareness in educators of the need to address individual student differences. Some revisionists have even attributed progress in public education to attempts at dealing with these individual differences (e.g., Tyack, 1974, p. 182).

In 1928, a survey of secondary school administrators addressed issues regarding individual differences among students. Billett (1932) concluded that homogeneous grouping, special classes, and the use of instructional plans incorporating unit assignments were the methods employed to individualize instruction. However, the amount of individualization evidenced varied between disciplines. A science education survey conducted at the time "indicated that the problem of adapting instruction to individual differences [had] received little attention" as evidenced by examination of courses of study (Beauchamp, 1932, p. 44).

From the 1930s until the mid-1950s little effort was directed toward curriculum development in science education. In the opinion of Baez (1976), science education at the secondary level during this 20-year span was typically unexciting. Baez felt that the sciences were taught in an authoritarian manner by ill-prepared teachers and that students were required to memorize sets of facts collected in textbooks written, for the most part, by secondary school science teachers. Occasionally teachers would demonstrate scientific phenomena, but the excitement of discovery was frequently missing (Baez, 1976).

After World War II, university scientists became aware that secondary school science textbooks were not abreast of scientific and technological developments. The Soviet launching of Sputnik in 1957 increased public concern regarding science education. The combined concern of scientists and the public resulted in a large scale reform movement in science education. The National Science Foundation (NSF) began to fund curriculum development projects and associated activities. Millions of dollars were spent on course improvement, workshops, teacher seminars and numerous professional development activities (Baez, 1976).

Curriculum development in the post-Sputnik years in science education became a popular and well-staffed undertaking. Its participants included not only teachers and scientists, but also administrators, educational scholars, and psychologists. Curriculum development findings were incorporated into elementary schools, secondary schools, and colleges. The presence of psychologists in the curriculum projects precipitated the inclusion of the philosophy of individualized instruction in many of the curricular projects.

Individualized instruction has become a frequently used term in American education. Individualized instruction is not only a "method of instruction" per se, but also a philosophy which permeates the teaching-learning environment. This philosophy has been instrumental in the development of a number of alternative methods of instruction.

A science educator today has available a number of instructional methods from which to choose. Some methods, such as the traditional lecture-laboratory, usually lack individualization. Other methods such as Keller's Personalized System of Instruction can be very individualized. Some methods are teacher-centered, whereas other methods are more student-centered. Some methods rely heavily on written programmed materials, some on audiovisual media, and others on interactive computer facilities.

Which method of instruction is the most effective? There is no simple answer to this question. Each method of instruction has its advocates and detractors. There is no single instructional method supported by a majority of high school and college faculty members.

Statement of the Problem

During the past fifty years many educational researchers have investigated the effectiveness of alternative methods of instruction in science. These investigations have resulted in the publication of a myriad studies. Conclusions derived from these investigations have frequently been contradictory. For example, a number of studies have been conducted comparing audio-tutorial and traditional instruction in general biology college courses. Some of these studies conclude that audio-tutorial instruction is more effective than traditional instruction (Meleca, 1970; Russell, 1968). Other studies indicate that traditional instruction is more effective than audio-tutorial instruction (Seal, 1976; VanderWal, 1973). In contrast, a few comparative studies conclude that audio-tutorial and traditional instruction are equally effective in the teaching of general biology (Belzer & Conti, 1973; Grobe, 1971 Weaver, 1969;).

Decades of research have shown that there is no consistently significant difference between one method of instruction and another. A review of the literature on alternative methods of instruction in science will reveal findings proclaiming the "superiority" of each instructional method available. Since teachers conventionally declare that teaching makes a difference in regards to student learning, researchers continue to design experiments to demonstrate this difference. As Taveggia (1977) illustrates there are numerous criticisms attributed to the research showing no differences between instructional methodologies. Yet, as Taveggia states, these criticisms "serve to 'bolster up' faith in conventional pedagogical dogma, and at the same time, provide justification for additional studies designed to demonstrate the superiority of one or another teaching method" (Taveggia, 1977, p. 120-121).

What is needed at this point is not another experiment, but a synthesis and integration of existing research. This synthesis should attempt to deal with the variability inherent in the existing studies, and as Light (1979) has noted, "try to harness the information contained in that variation" (p. 7). Controlling given variables may indicate that there is no best method of instruction, or it may indicate the superiority of one method of instruction under given circumstances.

Research studies comparing alternative methods of instruction have, of course been previously reviewed and integrated. These reviews, however, have usually been narrative in format and multidisciplinary in content. Meta-analysis, the statistical integration and analysis of research studies, has been applied to studies comparing alternative methods of instruction. Dubin and Taveggia (1968) analyzed studies comparing lectures, discussions and independent study. They found no significant differences in achievement among these three teaching methods. Kulik and his colleagues analyzed a number of studies comparing various methods of individualized instruction with college lecture classes and found some differences in achievement (Kulik, Cohen & Ebeling, 1979; Kulik, Kulik & Cohen, 1979a, 1979b, 1979c). None of these meta-analyses specifically dealt with alternative methods of instruction in science.

Although there are numerous studies involving alternative methods of instruction in science, these studies have not been quantitatively reviewed and synthesized. A meta-analysis would be of value to science educators by bringing order to a large collection of often contradictory research experiments. As Hurd (1971) noted, "we have reached a saturation point in several areas of science education research and the time has come for meaningful synthesis" (p. 245).

Purpose of the Study

The purpose of this study is to statistically integrate research studies conducted on alternative methods of individualized instruction in secondary school and collegiate science courses. Studies were analyzed to determine the relationship between instructional variables, study characteristics, and achievement and other outcomes. The alternative methods of individualized instruction included in this study were audio-tutorial instruction, computer-assisted instruction, Keller's personalized system of instruction and programmed instruction. Research studies incorporating unique individualized instructional systems were also included.

The two major goals of this meta-analysis were similiar to those goals described by Kulik, Kulik and Cohen (1979a). The first goal was to reach a conclusion about the experimental effect of individualized instruction in science; the second goal was to explain, if possible, the variation in any observed differences in effectiveness.

Hypotheses

Given that the purpose of education is to "prepare individuals for future responsiblities and for success in life, by means of acquisition of the organized bodies of information and prepared forms of skill" (Dewey, 1938, p. 18), and given that students approach this task of education with different skills and at different rates, it is a necessary derivative that instructional methodologies that respond to these differences will result in higher levels of student achievement. Based on this assumption, the author expected to find that individualized instruction in science was more effective than traditional instruction. The increased effectiveness of individualized instruction, however, may be due to variables such as mastery orientation or the degree of individualization relative to pacing, and not to the method of instruction per se. Such possibilities were considered in this study.

The effectiveness of instruction, within the context of this study, was measured by "effect size," which is a standardized difference between the group means of individualized and traditionally taught classes.

The following research hypotheses were generated from the above thesis:

<u>Overall Hypothesis</u>

Hypothesis one: The mean effect size of individualized instruction will be positive.

<u>Instructional Variable Hypotheses</u>

Hypotheses two through six: The variance of the effect size in individualized instruction is not accounted for by: (2) method of instruction, (3) mastery orientation, (4) degree of individualization relative to pacing, (5) degree of individualization relative to student selection of instructional delivery system, and, (6) degree of individualization relative to student initiated testing.

Although the instructional variables in the above hypotheses are all elements of individualization, there clearly are additional variables related to individualization which are not a part of this study. The author believes that the above variables will not explain all of the variance in effect size. Thus, the above hypotheses were stated in the null.

Design Variable Hypotheses

Hypotheses seven through twelve: The variance of the effect size in individualized instruction is not accounted for by: (7) meta-analyst's subjective rating of the study, (8) instrument development, (9) self-selection of treatment, (10) equivalence of subjects in the study, (11) historical effect, and (12) instructor continuity.

Miscellaneous Variable Hypotheses

Hypotheses thirteen through eighteen: The variance of the effect size in individualized instruction is not accounted for by: (13) source of publication, (14) year of publication, (15) instructional setting relative to grade level, (16) level of instruction, (17) subject matter taught, and (18) nature of instruction, i.e. replacement or supplement.

Summary Hypothesis

Hypothesis nineteen: The variance of the effect size in individualized instruction is not accounted for by the variance in instructional, design and miscellaneous variables.

Definitions

Audio-tutorial instruction (A-T) - An instructional method developed by Postlethwait (1963) using "modular independent study materials and a recorded audio component" (Fisher, 1976, p. 691). Audio-tutorial instruction involves the following components:

General Assembly Session (GAS) - A weekly meeting of the entire class for motivational lectures, films or examinations.

Small Assembly Session (SAS) - A weekly meeting of a small group of students for the purpose of discussion and quizzing.

Independent Study Session (ISS) - This principal activity in audio-tutorial instruction occurs in a learning center equipped with study guides, laboratory materials, audio tapes and visual aids. Students work individually at their own pace.

Computer-assisted instruction (CAI) - Instruction where the computer is used "to provide course content instruction in the form of simulations, games, tutorials and drill and practice" (Chambers & Sprecher, 1980, p.332).

Effect size (ES) - A standardized measure of the effectiveness of individualized instruction compared to traditional instruction. The effectiveness of the instruction is the difference in means between the instructional methods divided by an appropriate measure of the standard deviation.

$$ES = \frac{\text{mean of individually taught group} - \text{mean of traditionally taught group}}{\text{standard deviation}}$$

A positive effect size would reflect a greater effectiveness of individualized instruction, whereas a negative effect size would indicate a greater effectiveness of traditional instruction.

Individualized instruction - Instruction designed to accommodate student differences. The nature and degree of accommodation can vary. For example, one instructional method may vary the pacing of instruction, whereas another method of individualized instruction may allow students to select the method in which instruction is presented (i.e. film, lecture, reading, etc.). There are a number of methods of individualized instruction described in the literature. The nature and degree of individualization varies between and within each instructional method. However, there are some similarities which allow for categorization. In this study the following methods of individualized instruction were analyzed: audio-tutorial instruction, computer-assisted instruction, personalized system of instruction and programmed instruction.

Outcome - The dependent variable examined in a research study. In this meta-analysis the following outcomes were considered: achievement, attitude, retention of instructional content, and performance in subsequent courses.

Personalized system of instruction (PSI) - This method of instruction was first described by Keller (1968) and has frequently been called the Keller Plan or PSI. PSI usually involves the following components: 1) printed study guides, 2) mastery orientation, 3) student proctors, 4) self-pacing, and 5) occasional lectures for motivation.

Programmed instruction (PI) - This method of instruction involves the use of a text for the presentation of material in a step-by-step sequential manner. Programming is a procedure employed in many methods of individualized instruction. For the purpose of this study, programmed instruction is the presentation of information in a written format.

Science - A discipline utilizing the scientific method. The meta-analysis reported in this manuscipt included biology, chemistry and physics as the areas of science to be incorporated into this study of the effects of individualized instruction.

Traditional instruction - Instruction conducted in a face-to-face lecture format within a classroom. Implied in this is one teacher and a multiplicity of students. This method of instruction is usually teacher centered and group-paced with very little individualization of instruction.

Limitations

Since this study was a meta-analysis, it was limited by the accessibility and completeness of previously reported research. An attempt was made to include all relevant research published in journals and reported in doctoral dissertations. It was helpful to have access to the Library of Congress for use of the University Microfilms of doctoral dissertations. Bibliographic references to published or unpublished reports were utilized but an intensive search for unpublished reports and fugitive documents was not undertaken. Like many studies based on historical documents, this one thus suffers an unknown bias caused by differential representation of studies in the historical record. From what is known about the process of selection in professional journals, the author expects to find fewer published articles that found no statistically significant differences than may in fact have existed. It is the author's belief, however, that doctoral dissertation research is often reported even with statistically insignificant findings. Both kinds of documents were included in this meta-analysis and the results from these sources were compared.

CHAPTER 2

REVIEW OF THE LITERATURE

The literature was reviewed in two general areas: literature on research integration and literature on individualized instruction.

Research Integration

The assessment and integration of research findings has been a serious endeavor for a number of researchers. There are a number of procedures which have been recommended and utilized in the integration process. In this section of the literature review, a survey of a few of these procedures will be undertaken.

The literature review is a frequently encountered integrative activity. Most literature reviews in research have taken the form of a narrative description of existing research within a discipline. Literature reviews are an honest and valuable undertaking, but have their shortcomings. In reviewing a few studies, narratives are effective. However, as the number of studies increases it becomes difficult to assimilate and integrate information presented in a narrative format. Some research reviewers have dealt with this shortcoming by attempting to quantify research findings.

One method of quantifying research findings is to use a "voting method" (Light and Smith, 1971). In "taking a vote" the reviewer tallies the number of positively significant studies, negatively significant studies, and studies showing no significant differences. Light and Smith (1971) noted three disadvantages of the voting review. The first is that the "vote" is biased in favor of large sample studies. As sample sizes increase, the critical value required for significance decreases. This results in the chance occurrence of a greater number of significant large sample studies than of small sample studies. A second disadvantage of the voting review is that data not included in the voting process are excluded. The third disadvantage is that "voting" cannot account for interactions evident within the research studies. Glass (1976) noted an additional disadvantage to the voting review in that the review does not indicate the magnitude of differences being investigated. Two studies can each be significant at the .05 level, yet information about the relative size differences is lost.

Because of shortcomings in the voting review, Light and Smith (1971) recommended the use of a "cluster approach" for research integration. This approach involves subdividing studies into the smallest natural units of the educational process. These clusters are analyzed and differences controlled by analysis of covariance. Appropriate data are then combined, and the accumulated evidence from numerous studies integrated. Light and Smith (1971) summarized their procedure by stating that the "cluster approach enables us to pool data from conflicting studies by resolving their contradictions into representative differences among clusters" (p. 470).

The "cluster approach" to research integration allows the integration of similiar studies. In order for studies to be integrated the following standards must be met: 1) all subjects must be selected from a known population, 2) the variables must be identically measured in each study, and 3) the experimental design in all studies must be of equal quality.

The use of the "cluster approach" for research integration, however, is restricted because of its conservative requirements. It cannot be used in the integration of a large number of diverse studies. Moreover, it requires access to the original data.

A number of methods for combining the statistics and probabilities obtained from independent studies were evaluated in a paper by Rosenthal (1978). The methods he surveyed included adding probabilities, adding logs of probabilities, adding t-values, adding Z-scores, adding weighted Z's, testing the mean probability, testing the mean Z, counting, and blocking. Rosenthal noted that the method of choice depends upon the circumstances, but the "most serviceable under the largest range of conditions [is] the method of adding Z's" (p. 191).

Methods of adding statistics and probabilities are ways of attempting to integrate a large number of studies in a more sophisticated manner than the voting review. Nevertheless, the procedure is limited in that it does not quantify differences upon which the statistics and probabilities are based.

In 1976, Glass coined the term "meta-analysis" to identify "the statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings" (Glass, 1976, p.3). This type of analysis is of value in research integration. "By applying . . . the same objective methods that researchers use in analyzing results from an individual study, the meta-analyst is able to draw reliable, reproducible and general conclusions" (Kulik, Kulik, & Cohen, 1979a, p. 317).

A number of meta-analyses have been conducted in an attempt to answer a variety of questions. Hall (1978) conducted a meta-analysis of 75 studies reporting sex differences in decoding nonverbal cues. She found a significant difference in favor of females in the ability to decode nonverbal communication.

Smith and Glass (1977) conducted an extensive meta-analysis of 375 controlled evaluations of psychotherapy. They found a marked difference between treated and control groups and concluded that psychotherapy was effective.

Glass and Smith (1978) performed a meta-analysis on the relationship between school class-size and academic achievement. Eighty studies yielding 700 comparisons were collected and analyzed. Their results indicated that achievement dramatically increased when class size was reduced below 20 pupils. A similar study (Smith and Glass, 1979) investigated the relationship between class size and attitude. Fifty studies were analyzed and the results showed that decreased class size was associated with more positive attitudes.

A few meta-analyses have been published on comparisons of instructional methods. Dubin and Taveggia (1968) integrated studies conducted between 1924 and 1965 which compared college teaching methods. Differences in student performance between lecture, discussion, and independent study were quantified. The results of this integration led the authors to conclude that "the evidence from this analysis demonstrates that differences among teaching methods occur only rarely as measured by a final examination" (Dubin & Taveggia, 1968, p.8).

Kulik and his colleagues (Kulik, Cohen and Ebeling, 1979; Kulik, Kulik, & Cohen, 1979a, 1979b, 1979c) conducted several meta-analyses comparing alternative methods of individualized instruction (personalized system of instruction, audio-tutorial instruction, computer-assisted instruction, programmed instruction) and conventional lecture classes and found differences in achievement between the individualized instruction and traditional instruction.

Research on Individualized Instruction

There are numerous reviews of research on individualized instruction. Some of these reviews are extensive, and summarize research on several methods of individualized instruction, whereas others review the research on only one method of individualized instruction. In this chapter, reviews will be grouped according to the method of instruction. Reviews incorporating more than one instructional method are included in the first section.

Several Methods of Instruction

Seven reviews of research encompassing more than one method of instruction were read. Six of the seven reviews (Costin, 1972; Dubin & Taveggia, 1968; Jamison, Suppes &

Wells, 1974; Mallinson, 1976; Marchese, 1977; Welch, 1972;) concluded that all of the individualized instructional methods surveyed were equally effective. In contrast, McKeachie & Kulik (1975) indicated that the personalized system of instruction appeared to be the most effective method of instruction. Table 1 presents a summary of the views described in this section.

Dubin and Taveggia (1968) conducted a meta-analysis of 74 comparative studies of college teaching conducted between 1924 and 1965. They measured the effect of different instructional methods by calculating a standardized difference of group means on achievement measures (usually final exams). The instructional methods compared were lecture, discussion, supervised independent study and unsupervised independent study. The effects were compared using t-tests to determine if there were statistical differences. Dubin and Taveggia consistently found no differences between different instructional methods and stated that "the evidence from this analysis conclusively demonstrates that differences among teaching methods occur only rarely as measured by a final examination" (p. 8).

Costin (1972) reviewed research conducted between 1925 and 1967 comparing lectures with other instructional methods. He noted that lecturing was a common practice in

Table 1

Reviews of Research on Several Methods of Instruction

Reviewer	Year	Method of Instruction	N	Review Methodology	Conclusions
Dubin & Taveggia	1968	discussion independent study	74	meta-analysis	NSD
Costin	1972	discussion projects reading & PI	53	tabular	NSD
Jamison, Suppes & Wells	1974	PI, CAI, & others	100's	narrative tabular	NSD
McKeachie & Kulik	1975	PSI, PI, CAI	28	narrative	PSI most effective
Welch	1972	varied	30	voting	17 NSD 6 mixed 6 + exp 1 + control
Mallinson	1976	varied	40	narrative	inconclusive
Marchese	1977	varied	12	narrative	6 NSD 6 different

colleges and universities but its effectiveness was being questioned by many faculty members and administrators. Costin questioned the validity of previous extensive reviews of research by noting that less than one-third of the studies incorporated in his review were cited in previously published works.

The results of Costin's review of research were presented in a tabular fashion, with a limited amount of discussion regarding individual studies. About 20 studies comparing lecture and discussion methods showed similarities in achievement and seven studies comparing lecture and student-centered projects showed no consistent differences in achievement. Twenty-six studies comparing lecture with reading and use of self-instructional materials were reviewed. The self-instructional materials which were programmed in format tended to yield greater achievement. Costin (1972) concluded from his review that the "evidence fails to support popular derogation of the value of lectures in college and university teaching" (p. 26). There were no consistent differences in achievement between lecture and discussion, projects, reading or self-study. There was evidence that "programmed instruction may have an advantage over lectures for helping students learn facts and principles" (Costin, 1972, p. 22).

Jamison, Suppes and Wells (1974) conducted an extensive narrative review of the literature on alternative instructional media and methods (traditional classroom instruction, instructional radio, instructional television, programmed instruction, and computer-assisted instruction). They concluded that all media surveyed were effective for instruction. "Relatively few studies indicate a significant difference in one medium over another" (Jamison, Suppes & Wells, 1974, p. 55). The portions of the review on programmed instruction and computer-assisted instruction will be described in the section of this literature review designated for those instructional methods.

McKeachie and Kulik (1975) reviewed the literature dealing with effective college teaching and dealt with alternative methods of instruction in this context. They commented that in "recent years individualized instruction has become the fastest spreading innovation in higher education" (p. 166). The review was primarily devoted to a review of research on PSI with brief references to other methods of instruction. McKeachie and Kulik (1975) concluded that PSI is a very effective instructional method and that instructional television, programmed instruction, audio-tutorial instruction and computer-assisted instruction are "making a contribution to effective teaching, but none seems likely to revolutionize higher education" (p. 178).

Three reviews of research were specifically geared toward science; one summarized research published in 1968 and 1969 (Welch, 1972), another summarized research published in 1975 (Mallinson, 1976), and the third summarized research conducted between 1963 and 1974 (Marchese, 1977).

Welch (1972) reviewed a total of 30 studies involving comparisons of alternative methods of instruction. Seventeen studies concluded no significant differences, six had mixed results, six favored the experimental group and one favored the control group. Welch's review did not describe the types of alternative methods of instruction compared.

Mallinson's review included a summary of two reviews of research previously published. His conclusions were similar to other reviewers of literature on alternative methods of instruction. The findings of the 40 comparative studies reviewed were inconclusive. Mallinson (1976) noted that "a generalization that seems reasonable . . . is that achievement in any dimension depends greatly on factors other than the treatment" (p. 110).

Marchese (1977) selectively reviewed 12 studies comparing individualized and traditional modes of instruction in science. The studies selected were published between 1963 and 1974. Marchese's findings were similar to those of other reviewers. Six studies showed no significant differences in achievement between the alternative methods of instruction and six studies showed significant differences. PSI rated better than other forms of instruction.

Personalized System of Instruction (PSI)

In addition to McKeachie and Kulik's (1975) review, six other reviews of literature comparing Keller's personalized system of instruction (PSI) were reviewed. All of the reviewers concluded that PSI results in significantly greater achievement than traditional methods of instruction. Table 2 is a summary of the reviews of PSI.

Robin (1976) reviewed 39 studies comparing PSI with the lecture-discussion format at the college level. The studies were conducted between 1969 and 1975. Thirty comparisons resulted in significantly greater achievement for PSI groups, six comparisons indicated equal achievement and one study pointed toward superior performance in the lecture format. Two studies contained multiple comparisons showing mixed results. Four of the studies reviewed were in science (one in biology and three in physics).

Table 2

Reviews of Research on Personalized System of Instruction

Reviewer	Year	N	Review Methodology	Conclusions
Robin	1976	39	tabular	30 + PSI 6 NSD 2 mixed results 1 + traditional
Kulik, Kulik & Carmichael	1974	15	narrative	11 + PSI 4 NSD
Taveggia	1976	14	narrative	all favored PSI
Hursh	1976	16	narrative	favored PSI
Block & Burns	1976	24	voting, tabular	21 + PSI 3 + traditional
Johnson & Ruskin	1977	39	narrative	34 + PSI 5 NSD
McKeachie & Kulik	1975	15	tabular, narrative	11 + PSI 4 NSD
Kulik, Kulik Cohen	1979a	75	meta-analysis	ES = .49

Kulik, Kulik and Carmichael (1974) reviewed some studies on PSI in an article entitled "The Keller Plan in Science Teaching". Twenty-three studies were reviewed, of which eight were in the physical and natural sciences. Fifteen of the 23 studies cited compared examination performance of students in PSI and conventional courses. Four of the 15 studies indicated no significant differences; all of the remaining studies showed significantly greater achievment for PSI groups. Of the four science courses cited, two indicated no significant differences in achievement.

Taveggia (1976) summarized the results of 14 separate studies conducted between 1967 and 1974. One science study was included. The studies compared PSI and conventional teaching procedures by analyzing average student performance on course content examinations. All of the studies favored PSI. Taveggia concluded that "the Personalized System of Instruction has proven superior to the conventional teaching methods with which it has been compared" (p. 1029).

Hursh (1976) reviewed the literature comparing the effectiveness of PSI with other methods of instruction. Sixteen studies were reviewed using exam grades as course outcomes. Hurst concluded that students in PSI courses consistently had higher exam scores compared to students in courses taught by conventional methods of instruction.

Block and Burns (1976) located studies on PSI in journals and dissertations and reviewed those studies they felt were the "best of the mastery learning research to date" (p. 13). A total of 24 studies (two in science) were identified comparing PSI with a control group using achievement as the dependent variable. Twenty-one studies indicated greater achievement for PSI groups with 17 of the studies being statistically significant ($p < .05$). Three studies showed greater scores for the control groups at the same level of significance.

Johnson and Ruskin (1977) described various behavioral systems of instruction with a primary emphasis on PSI. The research reviewed in their document was summarized from a previous review by Kulik, Kulik and Smith (1976). The research summarized consisted of 39 experimental studies comparing PSI with lecture classes. Thirty-four studies showed a significant difference favoring PSI in end-of-course achievement, whereas the remaining five studies showed no significant differences. Johnson and Ruskin (1977) commented that all comparative studies indicate equal or better performance in PSI than conventional classes. "No reports have shown superior performance by conventionally taught students" (p. 73).

McKeachie and Kulik (1975) reviewed 15 studies on the comparative effectiveness of PSI and conventional instruction. Final examinations were used to measure effectiveness of each method of instruction. Eleven studies indicated that PSI was superior and four studies indicated no significant differences. McKeachie and Kulik (1975) commented that "content learning as measured by final examinations is adequate in Keller courses. In the published studies, final examination performance in Keller sections at least equals, and usually exceeds, performance in lecture sections" (p. 174).

Kulik, Kulik, and Cohen (1979a) applied Glass's meta-analysis procedure to the literature on Keller's PSI. A total of 75 courses taught conventionally and by PSI were incorporated in the analysis. The analysis consisted of t-tests for paired comparisons of means on outcome measures such as final exams, course grades, course ratings, course completion, and study time. The following results were observed: a) PSI was superior to conventional instruction as measured by final exams ($p < .0001$), b) there were significantly higher course grades in PSI ($p < .01$), c) student ratings were significantly higher in PSI courses and, d) there were no significant differences in course completion rates or in student study time between PSI and conventional classes.

Kulik, Kulik, and Cohen (1979a) used analysis of variance to further analyze the findings. They noted a significant interaction between method and discipline, with the highest outcomes occurring in psychology, math and engineering. There were no significant interactions between level of course, type of institution, method of subject assignment or the nature of the outcome analyzed.

Kulik, Kulik and Cohen's study was the first quantitative analysis of the literature on PSI. It's findings supported those of previous reviewers which extolled the virtues of PSI. The meta-analysis suggested that there was a difference in outcomes according to the discipline to which PSI was applied.

Audio-Tutorial Instruction (A-T)

Four reviews of literature on audio-tutorial instruction were read. Two reviews concluded that audio-tutorial instruction is superior to traditional instruction. Table 3 is a summary of the reviews on audio-tutorial instruction.

Mintzes (1975) reviewed 19 studies on A-T instruction in science. Seven of these studies were comparative in design. The results of these comparative studies indicated that, in most cases, A-T instruction was more effective than traditional instruction. Four studies

Table 3

Reviews of Research on Audio-Tutorial Instruction

Reviewer	Year	N	Review Methodology	Conclusions
Mintzes	1975	7	narrative	4 + AT 1 + traditional 2 NSD
Fisher & MacWhinney	1976	43	narrative	30 + AT 1 + traditional 13 NSD
Mallinson	1976	18	narrative	3 + AT 12 NSD
Kulik, Kulik & Cohen	1979c	48	meta-analysis	ES = .2

favored A-T, one study favored the conventional course, and two studies showed no significant differences.

Fisher and MacWhinney (1976) reviewed 89 studies on audio-tutorial instruction published between 1965 and 1975. Fisher (1976) summarized a selected segment of the more extensive review dealing with college level science. Both of the reviews involved some comparative studies between A-T and traditional instruction. There were 43 comparative studies in the Fisher and MacWhinney review. Student achievement was the dependent variable in each of the studies. A summary of the findings of these comparative studies indicated that 30 studies showed greater achievement in A-T instruction, 13 studies favored neither method of instruction, and one study showed greater achievement using the lecture format. Fisher (1976), in her science review, concluded that "A-T teaching may nearly always equal, and often be superior to, lecture methods with respect to content learning" (p. 695).

Mallinson (1976) reviewed studies on audio-tutorial instruction in science. Eight comparative studies conducted in secondary schools were included in the review. Six studies showed no significant differences, one was inconclusive and one study indicated that audio-tutorial instruction was superior to traditional instruction.

Similiar findings were observed in ten comparative studies conducted at the college level. Six showed no significant differences, two showed a significantly greater achievement in an experimental mixed format instructional setting, one study favored audio-tutorial instruction and one study found greater achievement in the lecture format.

Kulik, Kulik and Cohen (1979c) integrated 48 comparative studies on audio-tutorial instruction at the college level. Their analysis showed a small but positive effect (ES= .2) in favor of audio-tutorial instruction. Characteristics of studies such as research design, subjects taught, types of school and source of publication, did not appear to relate to the outcomes of the studies integrated.

Programmed Instruction (PI)

Programmed instruction is a form of individualized instruction which had its greatest activity in the 1960s. Some reviewers have tried to synthesize the results of research conducted on programmed instruction. Table 4 includes a summary of reviews on programmed instruction.

Jamison, Suppes and Wells (1974) summarized reviews totaling 200 studies comparing programmed instruction and traditional instruction. These summaries were derived from previously published reviews of literature. Jamison,

Table 4

Reviews of Research on Programmed Instruction and Computer-Assisted Instruction

Reviewer	Year	Method of Instruction	N	Review Methodology	Conclusions
Schramm	1964	PI	36	narrative	17 + PI 1 + traditional 18 NSD
Lange	1972	PI	112	narrative	41% + PI 10% + traditional 49% NSD
Hartley	1977	PI	40	meta-analysis	ES = .11
Kulik, Cohen & Ebeling	1979	PI	57	meta-analysis	ES = .24
Jamison, Suppes & Wells	1974	CAI	8	narrative	4 + CAI 4 + traditional
Kulik, Kulik & Cohen	1979b	CAI	59	meta-analysis	ES = .25

et. al. included Schramm's (1964) review which involved 36 studies comparing programmed with traditional instruction. Eighteen studies showed no significant difference in performance, 17 studies indicated greater achievement with programmed instruction, and one study showed greater achievement with the traditional lecture approach. Another review by Lange (1972) concluded that 41% of the studies favored programmed instruction, 49% found no difference between programmed instruction and traditional instruction, and 10% of the studies favored traditional instruction. Lange's review incorporated 112 studies.

Jamison, Suppes and Wells (1974) concluded their review of research on programmed instruction by stating that "programmed instruction is generally as effective as traditional instruction " (p. 41).

A meta-analysis was used by Hartley (1978) to integrate 40 studies on programmed instruction in elementary and secondary school mathematics. A total of 89 comparisons yielded results showing that programmed instruction was comparable to traditional instruction (ES = .11).

Kulik, Cohen and Ebeling (1979) conducted a meta-analysis on 57 comparative studies of programmed instruction at the college level. They found that programmed instruction had a small effect (ES = + .24) on student achievement. Study characteristics such as historical effects, instructor effects, subject area, type of school and level of instruction had little relationship with the outcomes of the studies. The publication date of a study did appear to be related to outcomes. More recent studies reported results having higher effect sizes than earlier studies.

Computer-Assisted Instruction (CAI)

Jamison, Suppes and Wells (1974) reviewed the research on computer-assisted instruction at the college level. They noted that there are limited studies on computer-assisted instruction and that no uniform conclusions can be made regarding the effectiveness of this medium. (See Table 4.)

Their review included eight comparative studies between CAI and traditional instruction. Four of the studies indicated that traditional instruction was more effective and four of the studies indicated that computer-assisted instruction was more effective. Jamison, Suppes and Wells (1974) concluded that "CAI is about as effective as traditional instruction when it is used as a replacement" (p. 55).

Kulik, Kulik and Cohen (1979b) synthesized the results of 59 evaluations of computer-assisted instruction at the college level. They found that computer assisted instruction had a small effect on achievement and attitude of the students toward instruction. The average effect size was found to be .25 for achievement and .24 for attitude. There appeared to be no relationship between the findings of the studies and characteristics of the studies such as design, settings or manner and date of publication.

Summary

This review does not include all reviews of literature on alternative methods of instruction. An attempt was made to make a representative selection of recent reviews. It is apparent that certain methods of instruction have been more extensively reviewed than others. This appears to be related to the quantity of research being conducted and to the interest of advocates of a given methodology in writing reviews of research.

The reviews were not independent. There was a good deal of overlap between the studies reviewed in different reviews of research on a given method. One comparison was made between two reviews and 50% of the references were duplicated.

A survey of the published reviews of research on individualized instruction indicates the presence of conflicting findings. A large number of reviewers support Dubin and Taveggia's (1968) conclusion that teaching method does not affect the outcome of instruction. However, just as there exist studies supporting the superiority of each instructional method, there likewise exist research reviews attesting to the superiority of different instructional methods. There is a tendency, however, to find a larger number of reviews proclaiming the superiority of the personalized system of instruction.

Based on the results derived from this review of research, the present meta-analysis should support one of the following conclusions: 1) all methods of individualizing instruction are equally effective or, 2) the personalized system of instruction is the most effective method of individualizing instruction.

CHAPTER 3

METHODOLOGY

This meta-analysis of alternative modes of teaching science courses consisted of a literature search for research findings comparing alternative methods of instruction in science, a subsequent review of this literature, and the selection of studies to be analyzed. The analysis involved a comparison of effect sizes and their relationship to other variables described in the research studies. The two major goals of the meta-analysis were similiar to those goals described by Kulik, Kulik and Cohen (1979a). The first goal was to reach a conclusion about the experimental effect in individualized instruction in science; the second goal was to explain the variation in the effect sizes obtained.

Literature Search

Documents incorporated in this study were identified in one of four ways. First a computer search of the Educational Resources Information Center (ERIC) was utilized with the key words: auto-instructional, individualized instruction, computer-assisted instruction, college science, secondary school science, chemistry, biology, and physics. Abstracts of each study so identified were scanned and appropriate studies identified. A total of 553 documents were initially identified, most of which were not comparative studies. Approximately 100 studies were selected for final review, and 34 were fully analyzed and incorporated into the meta-analysis.

Second, a computer search of Dissertation Abstracts was utilized on the key words: biology, chemistry, physics, individualized instruction, audiotutorial, autotutorial, computer assisted instruction, programmed instruction, Keller method and PSI. This search identified 73 dissertations most of which were not comparative studies. Twenty-three dissertations were read, 17 of which were incorporated in the study.

Third, the following reviews of literature involving alternative methods of instruction were used as sources of citations: Block and Burns (1976), Costin (1972), Dubin and Hedley (1969), Fisher (1976), Fisher and MacWhinney (1976), Hursh (1976), Jamison, Suppes, and Wells (1974), Johnson and Ruskin (1977), Kulik, Kulik, and Carmichael (1974), Kulik, Kulik, and Cohen (1979a), Mallinson (1976), Marchese (1977), McKeachie and Kulik (1975), Mintzes (1975), Razik (1971), Robin (1976), Taveggia (1976), and Welch (1972). Approximately 140 studies were identified from these reviews, of which about 50 were usable studies not identified with the computer searches.

Finally, bibliographies of studies located in the original search were scanned for research studies not identified by other sources. Approximately 20 studies were identified in this way.

Once documents were identified as potentially applicable, the following guidelines were used to determine which studies were to be included in the meta-analysis:

1. Only those studies comparing a traditional lecture method of science instruction with a method of individualized instruction were included.

2. Within a given study, multiple outcomes were incorporated in the meta-analysis. Some previous research integrators have avoided the use of more than one measurement per study (Dubin & Taveggia, 1968; Kulik, 1976). Dubin and Taveggia (1968) suggested that inclusion of multiple non-independent measures in an integration "obscures the 'true' value of the difference" (p. 62) since studies with multiple comparisons have a greater impact on final conclusions. To eliminate this problem Dubin and Taveggia (1968) included only one comparison of any two treatment groups. On the other hand, Glass (1978a) contends that much valuable information is discarded by eliminating multiple measurements. He feels that the central issue in decisions regarding the

inclusion of multiple outcomes revolves around the definition of the "unit of analysis" in a meta-analysis. Is the unit of analysis the "study" or the "finding"? If the "study" is considered the unit of analysis, findings are aggregated and many interesting relationships may be lost from the analysis. If the "finding" is considered the unit of analysis, a meta-analyst should deal with the interdependencies in a set of findings from a study. These interdependencies can be dealt with in two ways. First, an attempt can be made to assure that each finding measures a separate latent trait. Second, an inferential technique, the Tukey jackknife method, can be used to account for interdependencies in a large set of findings (Glass, 1978a).

Glass's (Smith & Glass, 1977; Glass & Smith, 1978) meta-analyses used "findings" as the unit of analysis. In his first meta-analysis on psychotherapy outcome studies, 833 measures were obtained from 375 studies. In Glass's (Glass & Smith, 1978) second meta-analysis involving the relationship of class-size and achievement, 700 measures were obtained from 80 studies.

For the meta-analysis of individualized instruction in science reported in this dissertation, multiple measures were incorporated from studies which included them. Thus, "findings" are the units of analysis.

Consideration was given to the interdependence of various outcome measures by attempting to exclude "findings" based on repeated measures of the same latent trait. For example, if a study reported midterm and final exam grades and an attitude measure, only the attitude measure and the final exam grade were used in the meta-analysis. It was assumed that the two exam grades were dependent and highly correlated. As a rule, the exam or grade measurement related to the longest treatment time was the outcome selected for inclusion. If a number of equally viable, but interdependent outcome measures were indicated, a random selection was made using a table of random numbers.

3. Poor quality of the research design did not necessarily eliminate a study from inclusion in the meta-analysis. Some researchers (Eysenik, 1978; Gallo, 1978; Mansfield, 1977) feel that only well-designed studies should be analyzed, because the incorporation of poorly designed studies into a meta-analysis leads to conclusions as weak as the inadequately designed studies included as units of analysis. However, Glass (1976) contends that elimination of weaker studies excludes a great deal of important and useful data. In addition, he feels that various "qualities" of design do not give very different findings. Smith and Glass's (1977)

meta-analysis of psychotherapy outcomes revealed that the " 'quality of design' accounted for one percent of the variance of 'effect size'" (Glass, 1978b, p. 3).

For the present study, an attempt was made to control for research design. A number of design features were identified and corresponding variables established. These variables are fully described in the data collection section of this chapter.

4. In order to be included in this meta-analysis, a study had to include measurements from which "effect sizes" could be calculated. A description of these procedures is included in the data analysis section of this chapter.

Data Collection

Each study was carefully read and the following information collected from each study:

1. Source of publication. This was categorized as a journal, document, dissertation, or book. If more than one source was available (for example, a journal article based on a dissertation), the one containing the most information was selected as the source of publication.

2. Year of publication. The year of completion of a doctoral dissertation was equated to year of publication.

3. Setting of the study. The type of educational

institution was categorized as either a secondary school, community college or four year college. A few studies were conducted at other educational institutions such as nursing schools. These were coded in a separate category of "other."

4. Subject matter. The subject matter was identified as a discipline of biology, chemistry or physics. A few geology, earth science and general science courses were reviewed and included in a category of "other."

5. Nature of the course. The course was classified as either introductory or advanced. An introductory course was defined as being general and basic in content. Advanced courses required the previous exposure of an introductory course.

6. Supplement or replacement of instructional method. Individualized instruction can either be a supplement to existing instructional techniques, or it can replace a previously existing technique and be an integral part of the instructional process. This dichotomy was used to classify each instructional method.

7. Method of instruction. The form of individualized instruction was categorized as either audio-tutorial, computer-assisted, personalized system of instruction, or programmed instruction. Instruction which was individualized but not clearly identified as one of the

previously identified methods, was included under a combination category.

8. Degree of individualization of the course relative to the pacing of instruction. This was a measure of the amount of individual pacing allowed in the course and was the most frequently employed technique for individualization. The degree of individualization relative to pacing was coded as follows: 1) no self pacing, 2) self-pacing on a daily basis, 3) self-pacing for short periods of time [weekly], 4) self-pacing for longer periods of time, and 5) self-pacing throughout the course.

9. Degree of individualization of the course relative to self-selection of delivery systems. This was a measure of the amount of freedom of choice that students had in the selection of the instructional delivery system. This variable was coded as follows: 1) no choice of delivery system, 2) choice of alternate delivery systems for supplementary instruction, 3) alternate delivery systems available for 1-50% of the course, and 4) alternate delivery systems available for 51-100% of the course.

10. Degree of individualization of the course relative to testing. The time for testing of content was either determined by the instructor or selected by the student. This variable measured the degree of freedom students had

in determining when they would be tested on material covered in the course.

11. Mastery orientation. Courses were categorized as to whether they were mastery oriented or non-mastery oriented.

12. Duration of instruction. This was a measure of the amount of time, in weeks, during which instruction was conducted using individualized instruction.

13. Number of students. The number of students in the experimental and the comparison groups were identified as separate variables.

14. Instructor effects. The individualized and traditional classes could either be taught by the same instructor or by different instructors. Which condition existed in each study was noted.

15. Historical effect. The individualized and traditional courses were either taught during the same semester or during different semesters.

16. Equivalence of subjects. Studies were categorized according to evidence of subject equivalence between the individualized and traditional groups. Equivalence was categorized as evidenced by the following: 1) randomization of the subjects, 2) categorical evidence on one to four measures, and 3) analysis of covariance with one, two or three covariates.

17. Self-selection of treatment. Students in the individualized classes either selected a method of individualized instruction or their course placement was determined by others.

18. Researcher involvement. This variable was meant to measure the personal involvement of the researcher in the results of the study. In some cases it was obvious that the researcher designed the content of the individualized instructional program, as well as evaluated its effects. These cases were categorized as having researcher involvement. In other cases, no indication was given that the researcher was personally involved in the outcome of the study.

19. Hawthorne effect. This is an effect which can occur because of the mere presence of an experimental condition. Studies were classified into two groups. One group included studies in which an attempt was made to control for the Hawthorne effect. The other group included studies where no apparent attempt was made to control for the Hawthorne effect.

20. Measurement. Three measurement variables were included in the meta-analysis. The development of a test was identified as being teacher developed, team developed or commercially available. The validity of a test was classified as either unreported or measured and reported.

The reliability of the test was included in the meta-analysis if it was described in the study.

21. Rating of the study. A subjective rating of the study was included as a variable in the meta-analysis. Studies were rated as being excellent, good, fair or poor. Ratings were assigned by the meta-analyst after carefully reading each study. The ratings were derived from a subjective evaluation of the quality of the research design coupled with the clarity of the presentation of the results.

22. Outcomes. The outcomes were the dependent variables encountered in the various studies. In this meta-analysis the outcomes were categorized as being either achievement, attitude, retention of subject matter, performanace in subsequent courses, tests on critical thinking, tests for logical reasoning or tests on the understanding of science. Outcomes relating to the retention of subject matter were further categorized on the amount of time between the end of treatment and the test for retention.

23. Effect sizes. Effect sizes were calculated using three different denominators. These denominators were the standard deviation of the traditionally taught group, the standard deviation of the group taught using individualized instruction, and the pooled standard deviation.

Data Analysis

Determination of Effect Size

To compare the effect of individualized instruction to traditional instruction it is important that a measure be used which quantifies the difference between instructional methods. Cohen (1969) proposed the use of an index "d" as a measure of effect size when comparing two population means:

$$d = \frac{m_A - m_B}{\sigma}$$

where: m = population mean of A or B

σ = standard deviation of either population A or B. The assumption is made that $\sigma_A = \sigma_B$

d = effect size index

The d-index has been incorporated by some meta-analysts for the measurement of effect sizes. It is, in a sense, a standardized score with a mean of zero and a unit standard deviation. Effect sizes calculated in this manner can be averaged and compared, allowing for ease in the quantification and comparison of numerous studies.

Glass (1978a) reintroduced the measurement of effect size by stating than an "informative and straightforward measure of experimental effect size" (p. 366) is:

$$ES = \frac{\overline{X}_E - \overline{X}_C}{s_x}$$

where: ES = effect size

$\overline{X}_E$ = mean of experimental group

$\overline{X}_C$ = mean of control group

s_x = standard deviation of both groups (assumed homogeneous)

When utilizing the above measure of effect size, difficulties are encountered if the group variances are heterogeneous. This difficulty is illustrated by the following example:

Experimental group: $\overline{X} = 25$, $s = 3$, $n = 25$

Control group: $\overline{X} = 20$, $s = 8$, $n = 25$

If used as basis standardization	Effect size is
s_E	1.67
s_C	.63
$(s_C + s_E) / 2$	.91

As illustrated, different effect sizes result depending upon which standard deviation is used. Should the population variance be estimated using the experimental group or control group standard deviation; or should a pooled estimate be used? There is no definitive answer to this question. Some meta-analysts have used the control group as the basis of standardization (Smith & Glass, 1977), whereas others have used the pooled deviations (Hall, 1978; Rosenthal, 1976).

Glass prefers to use the control group standard deviation because it provides an interpretation he finds substantively pleasing. If two methods are compared with a control situation, "standardization of mean differences by the control group standard deviation at least has the advantage of allotting equal effect sizes to equal means" (Glass, 1978a, p. 371). This position is reaffirmed by McGaw and Glass (1980), although they support the use of pooled standard deviations in certain situations (e.g., when the results of studies are reported using inferential statistics).

When comparing two methods of instruction there is no untreated control condition. Traditional instruction has been treated as the "control" by some meta-analysts (Athappilly, 1980; Hartley, 1978; Kulik, Kulik & Cohen, 1979a; Pflaum, Walberg, Karegianes & Rasher, 1980). The effect sizes reported by these researchers were calcuated using the standard deviations of traditionally taught groups.

There were many methods of individualized instruction investigated in this meta-analysis. Each study compared a method of instruction with traditional instruction. Since all studies used traditional instruction as a basis for comparison, the use of the traditional group standard deviation in calculating effect size may allow for more meaningful comparisons of effect sizes between instructional methods.

However, there is an advantage in using a pooled standard deviation in the calculation of effect sizes. Studies which did not indicate the required measurements for the calculation of effect sizes (i.e. $\overline{X}_E$, $\overline{X}_C$, and s) were included in the study and effect size measures approximated using procedures described by Glass (1978a). In situations where effect sizes were approximated the standard deviations were automatically pooled.

In order to qualify for approximation one of the following situations was required:

1. $\underline{t}$ given and an indication of sample sizes;

2. significance level of a mean difference, the direction of the effect, and the size of the samples; or

3. $\overline{X}_E$ and $\overline{X}_C$ plus the value of a multiple comparison statistic.

When results were presented in dichotomies or percentages, an attempt was made to recover underlying metric information. Effect size measures were obtained by determining the standard normal deviates corresponding to the percentages reported. The difference between these deviates was then used as an effect size measure.

Since many research studies did not report either the means or standard deviations, a large number of effect sizes were approximated. The use of the pooled standard deviation in calculating effect size resulted in a larger number of usable studies for incorporation into the meta-analysis.

There are advantages to using each method for calculating effect sizes. Both methods are correct; neither is false. Glass (1979) has suggested that both control group and pooled deviations be used in calculating effect sizes, and that the resulting effect sizes be correlated. If there is a high correlation between the two effect size measures, equivalent substantive conclusions will result regardless of which method is employed. A correlation of .99 between effect size measures has been reported elsewhere (Kulik, Kulik & Cohen, 1979c).

Statistics Employed

Analysis of the data involved: (a) descriptive statistics for the entire data set, (b) descriptive statistics for each of the individualized instructional methods, (c) a two-tailed t-test to test the overall effectiveness of individualized instruction in science,

(d) one-way analysis of variance to test the effect of selected variables on effect size, and (e) regression analyses in which effect sizes were regressed on selected independent variables.

Since numerous variables in the meta-analysis were assigned to categories of a nominal scale, dummy coding was used in the regression analysis. A set of dummy variables was created for k - 1 categories of each nominal variable. A score of 1 or 0 was then assigned to indicate the presence or absence, respectively, of membership in a category. Since dummy variables possessed a metric value they could be used in regression equations.

A limited number of t-tests and ANOVA's were calculated in order to limit the probability of making a Type I error. The level of significance selected was .05.

The Statistical Package for the Social Sciences (Nie, Hull, Jenkins, Steinbrenner & Brent, 1975) was the computer programming package used for the analyses.

CHAPTER 4

RESULTS

One hundred fifteen studies were included in this meta-analysis of individualized instruction in science. These studies were selected from an initial listing of about 800 studies. These selected studies provided a data set based on 26,004 students, who were approximately equally divided between traditional (N=13,824) and individualized (N=12,180) instructional groupings. Appendix A lists the studies included in this meta-analysis.

Forty-nine studies included more than one finding, and effect sizes were determined for each finding. The effect sizes were calculated using two different estimates for standardization; these were labeled ES_C and ES_P. The ES_C measure used the control group standard deviation, whereas the ES_P measure used a pooled standard deviation. There were 115 findings for which both ES_C and ES_P could be calculated; when they were correlated, the correlation coefficient of the effect size measurement was .904. A correlation this large indicates that over 80 percent of the variation in effect sizes can be explained with reference to the alternative form of calculating effect sizes. Since the substantive interpretations are unlikely to depend upon the method of calculating effect sizes, and because there were more comparisons available using effect sizes calculated with the pooled deviation, the pooled deviation measure of effect size was used in subsequent analyses.

All methods of individualized instruction were well represented in the meta-analysis. Achievement was the outcome most frequently studied. Table 5 shows the frequencies of instructional methods and outcome measurements from the 115 studies.

Descriptive statistics and multiple regression analyses were applied to all effect sizes (N=181). In addition, since 63 percent of the studies utilized achievement as a dependent variable, identical analyses were applied to the effect sizes of the achievement category (N=113).

Overall Hypothesis

Hypothesis one: The mean effect size of individualized instruction will be positive.

The primary goal of this study was to determine the effect of individualization in science instruction. The mean effect size, based on 181 findings reported in the

Table 5

Instructional Methods and Outcomes
Number of Studies and Effect Sizes

Variable	Number of Studies	Number of Effect Sizes
Method of instruction		
1. Audio-tutorial (A-T)	27	39
2. Computer-assisted (CAI)	11	13
3. Personalized system of instruction (PSI)	19	28
4. Programmed instruction (PI)	28	45
5. Combination	30	56
Total	115	181
Outcome measured		
1. Achievement	110	113
2. Attitude	22	27
3. Retention	19	22
4. Study time	1	1
5. Performance in subsequent courses	3	5
6. Others	9	13
Total	164	181

115 studies, was .352. Thus, the effect of exposure to individualized instruction in science was an increase of .35 standard deviations in measurable outcomes. The mean (50th percentile) for a class taught using individualized instruction would, therefore, be equal to the 64th percentile (+.35 standard deviation) of an otherwise equivalent class taught in a traditional lecture format. This effect size falls within the range of small effect sizes (0 to .5) according to Cohen (1969).

Hypothesis one stated that the effect size of individualized instruction would be positive. This hypothesis was tested for all outcomes using a two-tailed t-test; it was found that this effect size was significantly greater than zero ($t=8.517$, $p<.05$).

The mean effect size of the 113 findings reporting achievement was .353, again a small effect size (Cohen, 1969). A two-tailed t-test showed that this effect size was also significantly greater than zero ($t= 6.725$, $p<.05$).

These initial findings indicate that individualized instruction in science is more effective than traditional instruction. This conclusion may, however, be premature if the differences between two modes of instruction can be explained by factors other than differences in instruction. For example, if nearly all of the individualized treatments included self-initiated testing, and this factor explained the .35 standard deviation increase, then we should conclude that it was self-initiated testing that caused the observed differences in outcomes, not necessarily different modes of instruction per se.

This question was addressed in the second goal of this meta-analysis, i.e., to explain the variation in the obtained effect sizes. To meet this goal the variables encountered in the research studies were categorized as either instructional, design, or miscellaneous. Regression analyses were performed and the results grouped according to hypotheses relating to the variable categories.

Instructional Variable Hypotheses

Hypothesis two: The variance of the effect size in individualized instruction is not accounted for by the variance in the method of instruction utilized.

Tables 6 and 7 depict the effect sizes for the different instructional methods for all outcomes and for findings reporting achievement. The achievement-reporting studies classified as combination and computer-assisted had the largest effect sizes. Audio-tutorial instruction had the smallest effect size. The .56 effect size of computer-assisted instruction on achievement was a medium

Table 6

Effect Sizes of Instructional Methods for All Outcomes

Instructional Method	Effect Size	Number of Studies	Number of Effect Sizes
Audio-tutorial	.21	27	39
Computer assisted	.42	11	13
Personalized system of instruction	.42	19	28
Programmed instruction	.27	28	45
Combination	.47	30	56
Total	.35	115	181

$F(4,176) = 1.752$; Prob. = .859

Table 7

Effect Sizes of Instructional Methods for Achievement

Instructional Method	Effect Size	Number of Studies	Number of Effect Sizes
Audio-tutorial	.20	27	27
Computer assisted	.56	11	11
Personalized system of instruction	.34	16	17
Programmed instruction	.36	25	26
Combination	.41	30	32
Total	.35	109	113

$F(4,108) = .988$; Prob. = .583

size effect (.5 to .8) as defined by Cohen (1969). These differences were not, however, significant at the .05 level.

As part of this analysis, the effect size was regressed on a set of four dummy variables measuring membership in one of the five instructional categories. The coefficients of determination for these regressions were .038 for all outcomes and .035 for achievement. Thus, less than 4 percent of the variance in effect size was explained by the variance in method of instruction. Tables A and B in Appendix B illustrate the zero-order correlations among effect size and methods of instruction.

These findings indicate that the alternative methods of individualized instruction, per se, do not account for differences in outcome between individualized and traditional instruction in science. Since the effect sizes of individualized instruction are significantly greater than zero, one may still hypothesize that instructional variables _within_ the instructional methods may account for the variance in effect size. Some of these instructional variables were examined with hypotheses three through six.

Hypotheses three through six: The variance of the effect size in individualized instruction is not accounted for by: (3) mastery orientation, (4) degree of individualization relative to pacing, (5) degree of individualization relative to student selection of instructional delivery system, and, (6) degree of individualization relative to student initiated testing.

Table 8 lists the instructional variables and their coding categories. Comparisons were made of the effect sizes derived for each instructional variable investigated in the meta-analysis for each instructional method. These data are shown in Tables 9 and 10.

Examination of these tables may lead one to conclude that increased student choice and flexibility result in increased effectiveness as measured by effect size. The self pacing of instruction for an entire course (ES = .46) results in a higher effect size than daily pacing (ES = .26). Self initiated testing (ES = .46) leads to a higher effect size than the absence of this flexibility (ES = .29). Allowing students a choice of instructional delivery system (ES = .53) results in a higher effect size than having a uniform delivery system (ES = .30). In addition to choice and flexibility, a course requirement for mastery (ES = .53) results in a higher effect size than when this requirement is absent (ES = .28). These effect size differences are, for the most part, consistent for all methods of instruction.

Multiple regression analysis was used to identify the

Table 8

Instructional Variables
Number of Studies and Effect Sizes

Variable	Number of Studies	Number of Effect Sizes
Mastery orientation		
1. Non-mastery	83	129
2. Mastery	28	43
Self-initiated testing		
1. Present	38	65
2. Not present	74	112
Self-pacing of instruction		
1. Daily	19	28
2. Longer	53	78
3. Entire course	41	68
Choice of instructional delivery systems		
1. No choice	89	138
2. Choice	26	43

Table 9

Effect Sizes for Instructional Methods by Instructional Variables for All Outcomes

	TOTAL			AT			CAI			PSI			PI			COMB		
VARIABLES	ES	S.D.	n	ES	S.D.	n	ES	S.D.	n	ES	S.D.	n	ES	S.D.	n	ES	S.D.	n
Mastery orientation																		
Nonmastery	.28	.48	129	.21	.36	38	.26	.73	10	.22	.51	2	.29	.42	41	.36	.56	38
Mastery	.53	.63	43	.09	0	1	1.06	.73	2	.43	.46	26	.52	0	1	.67	.91	13
Self-pacing of instruction																		
Daily	.26	.43	28	-.38	0	1	.46	.85	4	-	-	-	.24	.34	19	.28	.08	4
Longer	.33	.46	79	.23	.36	37	.35	.33	3	.77	.38	2	.39	.51	18	.41	.58	19
Entire course	.46	.69	68	.08	0	1	.45	1.00	5	.39	.46	26	.36	.12	4	.55	.84	32
Self-initiated testing																		
Yes	.46	.68	65	-	-	-	-.43	.01	2	.41	.47	27	.48	.24	6	.56	.87	30
No	.29	.46	112	.21	.36	39	.57	.69	11	.50	0	1	.23	.44	38	.37	.50	23
Choice of delivery system																		
No choice	.30	.47	138	.19	.35	38	.39	.79	11	.39	.45	27	.25	.44	41	.41	.51	21
Choice	.53	.75	43	.74	0	1	.58	.30	2	1.03	0	1	.43	.23	4	.51	.83	35

Table 10

Effect Sizes for Instructional Methods by Instructional Variables for Achievement

	TOTAL			AT			CAI			PSI			PI			COMB		
VARIABLES	ES	S.D.	n	ES	S.D.	n	ES	S.D.	n	ES	S.D.	n	ES	S.D.	n	ES	S.D.	n
Mastery orientation																		
Nonmastery	.29	.47	82	.20	.39	26	.41	.74	8	.22	.51	2	.39	.38	25	.26	.54	21
Mastery	.50	.64	26	.09	0	1	1.06	.73	2	.36	.29	15	-	-	-	.67	1.02	8
Self-pacing of instruction																		
Daily	.35	.49	17	-.38	0	1	.70	.87	3	-	-	-	.35	.37	11	.24	.03	2
Longer	.32	.44	51	.22	.38	25	.35	.33	3	.50	0	1	.45	.45	11	.36	.58	11
Entire course	.42	.71	42	.08	0	1	.67	1.00	4	.33	.31	16	.30	.04	3	.48	.96	18
Self-initiated testing																		
Yes	.38	.72	38	-	-	-	-.43	.01	2	.33	.31	16	.56	.33	3	.44	.99	18
No	.33	.46	72	.20	.38	27	.65	.67	10	.50	0	1	.33	.41	22	.35	.46	12
Choice of delivery system																		
No choice	.30	.44	86	.18	.38	26	.55	.79	9	.34	.31	17	.36	.41	24	.19	.38	10
Choice	.52	.82	27	.74	0	1	.58	.30	2	-	-	-	.39	.23	2	.51	.91	22

instructional variables which account for variance in effect size. The results of this analysis are presented in Table 11 for all outcomes and Table 12 for achievement outcomes. The findings suggest that a small percentage of the variance in effect size is accounted for by the variance in instructional variables (R^2 = .065 for all outcomes and .073 for achievement). The zero-order correlations for these analyses are shown in Tables C and D in Appendix B.

Mastery orientation was the only instructional variable with a regression coefficient significantly greater than zero (see Table 12). Studies where individualized instruction required mastery exhibited effect sizes about a third larger than studies with no mastery requirement.

The evidence strongly indicates that the variance in effect size of individualized instruction in science is explained neither by method of individualized instruction nor by instructional variables such as mastery orientation, self pacing of instruction, student selection of instructional delivery system, or self-initiated testing.

Design Variable Hypotheses

Hypotheses seven through twelve: The variance of the effect size in individualized instruction is not accounted

Table 11

Regression Results of Instructional Variables for All Outcomes

Variable	B	Standard Error of B
Mastery	.305*	.121
Self-testing	.020	.120
No choice of delivery system	-.194	.102
Pace = daily	.088	.148
Pace = weekly	.063	.110
Constant	.344	

*$p < .05$

R^2 = .065; F(5,161) = 2.231; Prob. = .946

Table 12

Regression Results of Instructional Variables for Achievement

Variables	B	Standard Error of B
Mastery	.344*	.147
Self-testing	.192	.147
No choice of delivery system	-.209	.129
Pace = daily	.111	.186
Pace = weekly	-.001	.140
Constant	.080	

*$p < .05$

R^2 = .073; F(5,98) = 1.536; Prob. = .814

for by: (7) meta-analyst's subjective rating of the study, (8) instrument development, (9) self-selection of treatment, (10) equivalence of subjects in study, (11) historical effect, and (12) instructor continuity.

The relationship of design variables to effect size was next investigated in order to attempt further to explain the variance in effect size of individualized instruction. Table 13 lists the design variables analyzed in this analysis and the coding categories of these variables. Tables 14 and 15 depict the effect sizes for each categorical design variable for all outcomes and for achievement. The variables self-selection of treatment, subjective rating by meta-analyst and equivalence of subjects contain coding categories with observable differences in effect sizes. However, there is a great deal of variation among instructional methods for some of these variables. If subjects were self-selected into treatment groups (ES = .49), the effect size was larger than if subjects were not self-selected (ES = .32). Studies rated by the meta-analyst as excellent (ES = 1.08) had a higher effect size than those coded as good (ES = .35), fair (ES = .39) or poor (ES = .28). The use of one measure for determing equivalence of subjects (ES = .08) resulted in a lower effect size than the use of statistical control (i.e., analysis of covariance) for

Table 13

Design Variables
Number of Studies and Effect Sizes

Variable	Number of Studies	Number of Effect Sizes
Self-selection of treatment		
1. Yes	22	34
2. No	87	138
Continuity of instructors		
1. Same instructors	51	76
2. Different instructors	54	85
Historical effect		
1. Same semesters	97	114
2. Different semesters	18	26
Subjective rating by meta-analyst		
1. Excellent	6	15
2. Good	35	58
3. Fair	42	64
4. Poor	30	41
Instrument development		
1. Teacher developed	59	79
2. Team developed	19	26
3. Commercially available	31	48
Equivalence of subjects		
1. Absent	21	30
2. One measure	18	30
3. 2-4+ measures	21	32
4. Randomization	32	52
5. Covariates	20	30
Reliability of instrument	49	73

Table 14

Effect Sizes for Instructional Methods by Design Variables for All Outcomes

	TOTAL			AT			CAI			PSI			PI			COMB		
VARIABLES	ES	S.D.	n	ES	S.D.	n	ES	S.D.	n	ES	S.D.	n	ES	S.D.	n	ES	S.D.	n
Self-select of treatment																		
Yes	.49	.59	34	.20	.40	11	1.07	.79	3	.38	.23	10	.60	.08	4	.84	1.03	6
No	.32	.54	139	.22	.35	27	.07	.44	9	.49	.57	15	.23	.43	40	.45	.68	48
Continuity of instructor																		
Same instructor	.33	.49	76	.15	.29	13	.39	.83	6	.43	.17	8	.44	.31	15	.32	.58	34
Different instructors	.33	.57	86	.24	.41	22	.39	.75	6	.50	.58	15	.15	.47	25	.53	.75	18
Historical effect																		
Same semesters	.35	.56	115	.23	.38	33	.46	.75	12	.41	.47	26	.24	.41	42	.50	.74	42
Different semesters	.36	.56	26	.09	.26	6	-.03	0	1	.53	.28	2	.72	.39	3	.40	.68	14
Subjective rating																		
Excellent	1.08	.98	4	-	-	-	.79	0	1	1.10	1.57	2	-	-	-	.82	.90	12
Good	.35	.55	58	.36	.39	17	-.16	.45	3	.35	.14	6	.27	.27	7	.41	.74	25
Fair	.39	.44	64	.10	.31	17	.44	.73	5	.30	.31	13	.39	.46	17	.38	.50	12
Poor	.28	.50	41	.06	.20	5	.88	1.01	3	.50	.34	7	.17	.44	19	.25	.56	7
Instrument development																		
Teacher developed	.36	.58	80	.23	.42	18	.65	.67	7	.63	.63	10	.28	.41	28	.38	.81	17
Team developed	.45	.52	26	.48	.35	5	.36	0	1	.27	.40	7	.20	.62	7	.95	.41	6
Commercial	.28	.62	48	.01	.17	9	-.43	.01	2	.15	.15	4	.40	.40	3	.41	.72	30
Equivalence of subjects																		
Absent	.39	.50	30	.00	.03	3	1.07	.79	3	.39	.34	10	.52	.42	8	.07	.50	6
One measure	.08	.38	30	.00	.03	3	.21	.06	2	.25	.26	7	.00	.52	10	-	-	-
2-4+ measures	.34	.46	32	.41	.39	5	.63	.83	3	.49	.34	5	.17	.35	9	.28	.50	10
Randomization	.35	.52	53	.29	.41	17	-.01	.49	4	.80	.95	4	.19	.26	11	.50	.56	17
Covariates	.63	.80	30	.06	0	1	.79	0	1	.17	.26	2	.60	.32	5	.70	.93	21

Table 15

Effect Sizes for Instructional Methods by Design Variables for Achievement

	TOTAL			AT			CAI			PSI			PI			COMB		
VARIABLES	ES	S.D.	n	ES	S.D.	n	ES	S.D.	n	ES	S.D.	n	ES	S.D.	n	ES	S.D.	n
Self-select of treatment																		
Yes	.55	.67	22	.26	.51	7	1.07	.79	3	.46	.23	6	.59	.05	2	.82	1.25	4
No	.31	.51	84	.19	.35	19	.19	.42	7	.32	.34	8	.34	.41	24	.39	.73	26
Continuity of instructor																		
Same instructor	.30	.46	50	.08	.24	9	.52	.86	5	.43	.18	7	.48	.35	10	.20	.50	19
Different instructors	.35	.57	51	.24	.44	16	.55	.72	5	.32	.39	7	.31	.43	13	.50	.92	10
Historical effect																		
Same semesters	.36	.56	95	.21	.41	22	.61	.72	10	.32	.31	15	.31	.38	23	.46	.81	25
Different semesters	.30	.54	18	.12	.28	5	-.03	0	1	.53	.28	2	.72	.39	3	.23	.75	7
Subjective rating																		
Excellent	.77	1.05	7	-	-	-	.79	0	1	-.01	0	1	-	-	-	.93	1.21	5
Good	.33	.61	35	.34	.43	12	-.04	.56	2	.33	.10	2	.25	.30	5	.40	.86	14
Fair	.29	.38	39	.08	.34	11	.62	.72	4	.26	.25	8	.53	.32	8	.20	.17	8
Poor	.36	.54	30	.09	.22	4	.88	1.01	3	.52	.37	6	.29	.48	12	.27	.67	5
Instrument development																		
Teacher developed	.35	.55	57	.17	.46	12	.80	.60	6	.45	.31	9	.36	.41	18	.22	.83	12
Team developed	.46	.46	19	.45	.40	4	.36	0	1	.09	.21	5	.38	.37	5	1.08	.34	4
Commercial	.26	.68	23	.04	.23	5	-.43	0	1	.26	0	1	.56	0	1	.36	.81	15
Equivalence of subjects																		
Absent	.41	.55	23	.02	.02	2	1.07	.79	3	.46	.34	8	.49	.51	5	.01	.53	5
One measure	.13	.33	15	.09	.30	7	-.17	0	1	.23	.38	3	.19	.45	4	-	-	-
2-4+ measures	.42	.50	19	.53	.31	4	.63	.83	3	.10	0	1	.26	.37	5	.43	.59	6
Randomization	.26	.39	32	.23	.48	11	.13	.50	3	.33	.18	3	.17	.27	7	.40	.43	8
Covariates	.55	.86	21	.06	0	1	.79	0	1	.17	.26	2	.72	.20	4	.58	1.08	13

determining equivalence (ES = .63).

Multiple regression analysis was used to identify design variables which account for variance in effect size. The results of this analysis are presented in Table 16 for all outcomes and Table 17 for achievement; the zero-order correlations for these analyses are shown in Tables E and F in Appendix B. The regression equation for all outcomes contains regression coefficients which are significantly greater than zero (self-selection of treatment, instrument development and equivalence of subjects). When only those studies involving achievement were analyzed, a regression coefficient significantly greater than zero was obtained for self-selection of treatment.

Studies involving self-selection of treatment exhibited effect sizes approximately one half standard deviation larger than studies that did not involve self-selection. The regression results for all outcomes showed that studies with teacher or team developed tests had effect sizes that were about one third larger than studies using commercially available tests. Studies where one measure was used to determine the equivalence of subjects reported effect sizes of about one half standard deviation less than studies using no measure, two or more measures or analysis of covariance.

Table 16

Regression Results of Design Variables for All Outcomes

Variable	B	Standard Error of B
Self-selection	-.457*	.157
Instructor continuity	-.004	.107
Historical effect	.109	.170
Rating of study		
Excellent	.486	.259
Good	.002	.158
Fair	-.027	.153
Instrument development		
Teacher developed	.233*	.119
Team developed	.322*	.146
Equivalence of subjects		
Absent	-.001	.195
One measure	-.466*	.159
2+ measures	-.057	.142
Covariates	-.132	.160
Constant	.991	

*$p < .05$

$R^2 = .200$; $F(12,115) = 2.423$; Prob. = .992

Table 17

Regression Results of Design Variables for Achievement

Variable	B	Standard Error of B
Self-selection	-.460*	.174
Instructor continuity	.009	.124
Historical effect	.007	.183
Rating of study		
Excellent	-.213	.342
Good	-.065	.180
Fair	-.163	.178
Instrument development		
Teacher developed	.197	.151
Team developed	.303	.173
Equivalence of subjects		
Absent	.112	.218
One measure	-.289	.193
2+ measures	.141	.172
Covariates	.087	.178
Constant	1.063	

*$p < .05$

$R^2 = .210$; $F(12,71) = 1.576$; Prob. = .881

These regression analyses of design variables yielded larger coefficients of determination (R^2 = .20 for all outcomes and .21 for achievement) than the regression analyses of instructional variables, and suggest that factors related to the design of studies have more of an influence on effect size than factors related to instruction. About 20 percent of the variation in effect sizes is apparently due to variations in the design of the studies; while most of the variation in effect sizes remains unexplained, at least a portion is artificially due to the methods of measurement.

The reliability of the instruments used in measuring outcomes was reported in some studies (N = 73 for all outcomes and 43 for achievement). The mean reliability coefficient was high ($\bar{X}$ = .84) with a small amount of variance reported (S.D. = .08 for all outcomes and .09 for achievement). Multiple regression analysis showed that less than 1 percent of variance in effect size was accounted for by the variance in the reliability of the instrument used in measuring outcome (R^2 = .000 for all outcomes and .005 for achievement).

Miscellaneous Variable Hypotheses

Hypotheses thirteen through eighteen: The variance of the effect size in individualized instruction is not accounted for by: (13) source of publication, (14) year of publication, (15) instructional setting relative to grade level, (16) level of instruction, (17) subject matter taught, and (18) nature of instruction, i.e., replacement or supplement.

Table 18 lists the miscellaneous variables and their coding categories. A majority of the courses analyzed were introductory. Individualized instruction was usually used as a replacement for existing instruction. About twice as many studies were reported in dissertations than in journal articles; the mean year of publication was 1972. Biology and chemistry were equally represented and comprised 67 percent of the sample. Four year collegiate institutions were the most frequent setting for the comparative studies.

Tables 19 and 20 illustrate effect sizes for the miscellaneous variables for each instructional method. Examination of these tables shows fairly uniform effect sizes. The only variables exhibiting sizable differences for coding categories are level of instruction for all outcomes and nature of instruction, level of instruction and source of information for the achievement outcomes. Higher effect sizes were observed for advanced courses (ES = .55) than introductory courses (ES = .32). Courses incorporating individualized instruction as a supplement

Table 18

Miscellaneous Variables
Number of Studies and Effect Sizes

Variable	Number of Studies	Number of Effect Sizes
Level of instruction		
1. Introductory	100	161
2. Advanced	14	19
Nature of instruction		
1. Replacement	101	166
2. Supplement	13	15
Source of publication		
1. Journal	38	52
2. Document	8	14
3. Dissertation	65	110
4. Book	4	5
Setting of study		
1. Secondary school	34	58
2. Community college	12	20
3. 4 yr institution	67	74
Subject of instruction		
1. Biology	40	59
2. Chemistry	37	58
3. Physics	20	37
4. Other	18	27
Year of publication		
1961	1	3
1963	4	5
1964	4	4
1965	3	5
1966	2	3
1967	2	3
1968	4	9
1969	8	9
1970	13	24
1971	11	17
1972	13	18
1973	12	19
1974	11	15
1975	11	21
1976	10	13
1977	5	10
1978	1	3

Table 19

Effect Sizes for Instructional Methods by Miscellaneous Variables for All Outcomes

	TOTAL			AT			CAI			PSI			PI			COMB		
VARIABLES	ES	S.D.	n	ES	S.D.	n	ES	S.D.	n	ES	S.D.	n	ES	S.D.	n	ES	S.D.	n
Level of instruction																		
Introductory	.32	.53	161	.16	.36	33	.47	.83	10	.44	.49	24	.26	.42	43	.40	.64	51
Advanced	.56	.71	19	.45	.30	6	.26	.28	3	.30	.16	4	-.18	0	1	1.24	1.12	5
Nature of instruction																		
Replacement	.35	.55	165	.21	.36	39	.05	.53	5	.42	.46	28	.28	.43	37	.47	.72	56
Supplement	.39	.65	15	-	-	-	.65	.78	8	-	-	-	.09	.32	7	-	-	-
Source of publication																		
Journal	.38	.51	52	.38	.44	16	.41	1.13	4	.47	.26	15	.13	.59	10	.54	.52	7
Document	.47	.82	14	.13	.15	3	.33	.76	5	-	-	-	.24	.02	2	1.02	1.24	4
Dissertation	.32	.55	110	.08	.26	20	.53	.25	4	-	-	-	.31	.38	33	.41	.69	45
Book	.40	.33	5	-	-	-	-	-	-	.40	.33	5	-	-	-	-	-	-
Setting of study																		
Secondary school	.25	.48	57	-.38	0	1	.08	.60	4	-	-	-	.24	.40	23	.30	.52	29
Community college	.40	.56	20	.18	.35	10	-	-	-	.89	1.17	3	.43	.25	4	.60	.57	3
4 year institution	.41	.60	100	.26	.36	26	.57	.76	9	.36	.30	25	.26	.50	18	.70	.93	22
Subject of instruction																		
Biology	.30	.44	59	.19	.37	30	.36	0	1	.64	.24	3	.21	.40	11	.54	.55	14
Chemistry	.36	.64	58	.24	.43	3	.52	.88	8	.63	.75	7	.25	.43	21	.32	.73	19
Physics	.41	.61	37	.27	.37	5	.23	.52	4	.35	.26	13	-.03	.48	3	.72	.89	12
Other	.37	.54	27	.37	0	1	-	-	-	.17	.36	5	.47	.41	10	.38	.72	11

Table 20

Effect Sizes for Instructional Methods by Miscellaneous Variables for Achievement

	TOTAL			AT			CAI			PSI			PI			COMB		
VARIABLES	ES	S.D.	n	ES	S.D.	n	ES	S.D.	n	ES	S.D.	n	ES	S.D.	n	ES	S.D.	n
Level of instruction																		
Introductory	.32	.52	98	.15	.37	23	.67	.81	8	.34	.33	15	.36	.38	24	.31	.67	28
Advanced	.55	.78	14	.46	.38	4	.26	.28	3	.33	.10	2	-.18	0	1	1.16	1.27	4
Nature of instruction																		
Replacement	.33	.55	100	.20	.38	27	.17	.52	4	.34	.31	17	.38	.41	20	.41	.79	32
Supplement	.53	.65	12	-	-	-	.77	.74	7	-	-	-	.20	.31	5	-	-	-
Source of publication																		
Journal	.37	.50	35	.33	.49	12	.69	1.21	3	.42	.27	9	.28	.48	6	.33	.47	5
Document	.51	.95	9	.07	.15	2	.48	.79	4	-	-	-	.22	0	1	1.16	1.99	2
Dissertation	.31	.54	65	.10	.26	13	.53	.25	4	.05	.20	4	.39	.39	19	.37	.75	25
Book	.47	.33	4	-	-	-	-	-	-	.47	.33	4	-	-	-	-	-	-
Setting of study																		
Secondary school	.27	.45	34	-.38	0	1	.24	.62	3	-	-	-	.32	.42	13	.28	.47	17
Community college	.36	.42	12	.23	.33	7	-	-	-	.23	.34	2	.49	.38	2	1.25	0	1
4 year institution	.41	.63	65	.24	.40	18	.67	.74	8	.36	.31	15	.39	.40	11	.54	1.10	13
Subject of instruction																		
Biology	.28	.43	40	.20	.41	21	.36	0	1	.64	.24	3	.20	.44	07	.42	.52	8
Chemistry	.41	.64	36	.37	.52	2	.66	.85	7	.35	.34	5	.37	.34	12	.33	.91	10
Physics	.39	.69	19	.02	.17	3	.39	.50	3	.28	.11	5	-.31	0	1	.74	1.02	7
Other	.35	.50	18	.37	0	1	-	-	-	.19	.41	4	.65	.28	6	.19	.66	7

(ES = .53) exhibited higher effect sizes than courses where the individualized instruction was a replacement to existing instruction (ES = .33). Studies reported in documents had a higher effect size (ES = .51) than studies published in dissertations (ES = .31), journal articles (ES= .37), or books (ES = .47).

Multiple regression analysis was used to identify the miscellaneous variables which account for the variance in effect size. The results of this analysis are presented in Table 21 for all outcomes and Table 22 for achievement; the zero-order correlations for these analyses are shown in Tables G and H in Appendix B.

None of the regression coefficients was significantly greater than zero. This analysis yielded small coefficients of determination (R^2 = .04 for all outcomes and .05 for achievement). Thus, less than 5 percent of the variance in effect size is explained by the variance in the categorical miscellaneous variables.

A regression analysis of year of publication indicated that this miscellaneous variable accounted for a very small amount of variance in effect size (R^2 = .01 for all outcomes and .001 for achievement).

Table 21

Regression Results of Miscellaneous Variables for All Outcomes

Variable	B	Standard Error of B
Level of instruction	.235	.149
Replacement of instruction	-.023	.160
Source of publication		
Journal	.054	.266
Document	.150	.302
Dissertation	.053	.265
Setting of study		
Secondary school	-.100	.110
Community college	.078	.151
Subject of instruction		
Biology	-.097	.142
Chemistry	-.007	.142
Physics	.016	.148
Constant	.108	

R^2 = .038; F(10,169) = .665; Prob. = .244

Table 22

Regression Results of Miscellaneous Variables for Achievement

Variable	B	Standard Error of B
Level of instruction	.218	.177
Replacement of instruction	.144	.186
Source of publication		
Journal	-.091	.305
Document	-.008	.351
Dissertation	-.116	.302
Setting of study		
Secondary school	-.067	.139
Community college	.062	.193
Subject of instruction		
Biology	-.084	.174
Chemistry	.048	.176
Physics	.021	.195
Constant	.062	

R^2 = .05; F(10,101) = .536; Prob. = .139

Summary Hypothesis

Hypothesis nineteen: The variance of the effect size of individualized instruction is not accounted for by the variance in instructional, design and miscellaneous variables.

All of the variables were entered into one large regression equation for the purpose of explaining as much of the variation in effect size as possible. Pairwise deletion of missing data was used in order to maximize the amount of information incorporated into the analysis.

The regression analysis yielded coefficients of determination of .33 for all outcomes and .32 for achievement. Thus, approximately 30 percent of the variance in effect size was explained by the variance in instructional, design and miscellaneous variables. These results clearly demonstrate that while about a third of the variation in effect sizes can be explained by variables that measure variation in research design, methods of individualized instruction, and other miscellaneous variables, most of the variation in effect sizes remains unexplained.

Conclusions

The findings of this meta-analysis suggest that individualized instruction in science is more effective than traditional instruction. This increased effectiveness is only minimally explained by a number of instructional, design and miscellaneous variables.

The importance of these results rests precisely with the inability to explain away the effect sizes. They were not due to instructional factors other than individualization. They were not due to factors relating to the design of research. They were not due to a variety of miscellaneous factors associated with these studies. They were not due to all of these grouped together. Therefore, the author concludes that they must be due to individualization itself. The implications of this conclusion are considered in the next chapter.

CHAPTER 5

DISCUSSION AND CONCLUSIONS

The purpose of this study was to quantitatively integrate existing research that investigated alternative methods of individualized instruction in secondary school and collegiate science courses. The major goals of the meta-analysis were to determine the experimental effect of individualized instruction in science, and to explain the variation in observed differences in effectiveness to the extent possible.

The study included an analysis of 115 studies comparing five methods of individualized instruction in science with traditional lecture instruction. The effectiveness of individualized instruction was measured in terms of effect size, a standardized difference in means between the individualized and the lecture classes.

Instructional, design and miscellaneous variables in the research studies were identified and multiple regression analysis was used to attempt to explain the variation in effect size.

Discussion

The mean effect size of the 181 findings integrated in this meta-analysis was .35; that is, the mean of the students taught with individualized instruction was .35 standard deviations above the mean of students taught with traditional instruction. Even though this was a small effect size (see Cohen, 1969), it was significantly greater than zero. If one assumes that the studies incorporated in this meta-analysis represent the universe of individualized and traditional methods of instruction, we can conclude that individualized instruction is somewhat more effective than traditional instruction, whether one considers achievement alone, or achievement combined with several other outcome measures.

The observed increase in effectiveness of individualized instruction over traditional instruction could be due to several possible conditions. First, these results could be biased because those studies showing an increase in effectiveness for the "experimental" method were more likely to be published. But note that a majority (57%) of the studies incorporated in this meta-analysis was dissertations. There is probably less pressure to "publish" significant results with dissertations than with journal articles. In addition the effect sizes reported for dissertations and journals were fairly similiar (ES = .32 for dissertations and .38 for journals). Thus, it seems unlikely that these conclusions are seriously biased by problems of selection.

A second condition which may explain the increased effect size of individualized instruction over traditional instruction is the impact of design and miscellaneous variables on the results of the comparative studies. Previous invesigators have suggested that the design of experiments influences effect size (Eysenik, 1978; Gallo, 1978; Mansfield, 1977). The implication is that poorly designed studies yield higher effect sizes than well designed studies. Glass (1976), on the other hand, feels that various "qualities" of design do not result in different meta-analytic findings.

The results of this meta-analysis indicated that approximately 20 percent of the variance in effect size was accounted for by the variance in design variables. This was a larger percentage than reported in Glass's meta-analysis of psychotherapy where the " 'quality of design' accounted for one percent of the variance of 'effect size'" (Glass, 1978b, p.3).

Self-selection of treatment, instrument development and equivalence of subjects were the design variables having the most influence on effect size. When subjects self-selected their treatment groups, the effect size was higher than when there was no self-selection. This finding corresponds with other results in this meta-analysis which show that giving student choice and flexibility leads to larger effect sizes.

Studies employing teacher or researcher developed instruments resulted in higher effect sizes than those using commercially available instruments. This finding may indicate that testing instruments developed by teachers or researchers are biased. Frequently a teacher initiates an individualized instructional program and then wishes to evaluate its effectiveness. Much time and effort have often been invested in the new system. The teacher may unconsciously develop a test which selectively measures the strongest segments of the new project, i.e., a biased instrument. It may also be true, however, that the locally developed instrument provides a more accurate reflection of the content covered in the instruction than a commercial instrument. The resulting higher effect size would, in this situation, not be from a biased instrument but from a more "finely tuned" speciality instrument.

The equivalence of subjects influenced effect size in an interesting way. It appears that as the design became stronger (i.e., the equivalence of subjects increased), the effect size increased. When subjects were equivalent on one measure, the effect size was .08. Two to four measures of equivalence resulted in an effect size of .34. When analysis of covariance was used for statistical control, the effect size was .63. These findings contradict the assumption that poorly designed studies yield higher effect sizes.

Kulik, Kulik and Cohen (1979a, 1979b) found that a continuity of instructors between the individualized and traditional classes influenced effect size. This influence was not, however, observed in this meta-analysis.

None of the miscellaneous variables appeared to influence effect size. Less than 5 percent of the variance in effect size was explained by variables such as instructional setting, subject matter, and publication.

Kulik and his colleagues found significant effects of subject matter (Kulik, Kulik & Cohen, 1979a), source of publication (Kulik, Kulik & Cohen, 1979c), and year of publication (Kulik, Cohen & Ebeling, 1979), on effect size. These effects were not observed in this meta-analysis.

If neither of the above conditions explain variation in effect sizes, a third condition which may explain the increased effect size of individualized instruction is a true difference in effectiveness between individualized and traditional methods of instruction, and these differences exist because of instructional variables.

The method of instruction used for individualization (e.g., CAI, PSI, A-T, etc.) may be the variable responsible for the increased effectiveness of individualized instruction. The differences in effect size measures between instructional methods were tested and found to be not significant. Less than 4 percent of the variance in effect size was explained by the variance in method of instruction. Thus, it appears that method of instruction is not the variable which accounts for the increased effectiveness of individualized instruction.

The effect sizes for various methods of instruction closely approximate those reported by Kulik and his associates. Kulik's meta-analyses synthesized research on individualized instruction at the college level. All types of subject matter were incorporated in his analyses. Kulik, Kulik and Cohen (1979c) reported an effect size of .2 from an analysis of 48 studies on audio-tutorial instruction. The present meta-analysis yielded an effect size of .21 for 27 audio-tutorial studies. Kulik, Kulik and Cohen (1979b) calculated an effect size of .25 for 59 studies of computer assisted instruction. This dissertation reports a higher effect size of .42 for 11 such studies. Within this meta-analysis, computer assisted instruction had a small sample size and a large variance in effect size measurements, which may explain the difference between the results of the two meta-analyses. In another study, Kulik, Cohen and Ebeling (1979) reported an effect size of .24 for 57 studies on programmed instruction. The present meta-analysis reports an effect size of .27 for 28 programmed instruction studies. Finally, Kulik, Kulik and Cohen (1979a) reported an effect size of .49 for 61 studies with the personalized system of instruction. This dissertation reports an effect size of .42 for 19 such studies.

Method of instruction, per se, did not apparently explain the variance in effect size. Other instructional variables such as mastery orientation or the pacing of instruction, may do so. Though less than 7 percent of the variance in effect size was explained by the variance in instructional variables there were differences in effect sizes for different instructional variables.

The positive effect of mastery orientation on effect size was a predictable finding. If an instructional process requires the mastery of subject matter, this logically leads to higher achievement measurements and subsequently higher effect sizes. Instructional variables which allow for increased student choice and flexibility resulted in increased effect sizes. If students are given a choice, they appear to choose that which leads to greater achievement.

Conclusions

The following conclusions were drawn based on the results of this meta-analysis of individualized instruction in science. First, individualized instruction was somewhat more effective than traditional instruction. Second, all five methods of individualized instruction analyzed, i.e., audio-tutorial, computer-assisted, personalized system of instruction, programmed instruction and a combination approach, were not differentially effective when compared to traditional instruction. Third, design and miscellaneous variables identified in this study contribute little to the variance in effect size, although design variables account for more variance than miscellaneous or instructional variables. Fourth, instructional variables identified in this study contribute little to the variance in effect size. Finally, the increased effectiveness of individualized instruction may be due to individualization, per se; in other words, the presentation of instruction to the individual student, and not to a group of students.

Implications

Individualized instruction appears to enhance learning. The overall gain in effectiveness over traditional instruction is approximately 14 percentile points. Although not a large gain, it is significantly greater than zero.

Since the present meta-analysis indicates that all five methods of individualized instruction in science are about equally effective, the decision of which method to employ should be based on other factors, such as cost and personal preference. Are there differences in the time, effort and expense required for the design and implementation of an established individualized instructional program (i.e., one of the five methods analyzed in this dissertation)? The time and effort required is probably similiar for most of the instructional methods. If any one method is more time consuming than the others, it is probably computer-assisted instruction. The equipment requirements, however, vary. Computer-assisted instruction and audio-tutorial instruction require special hardware for their implementation and are consequently more expensive to implement. It therefore appears that the personalized system of instruction, programmed instruction, or a combination approach are the most cost effective method of individualizing instruction.

Should science teachers eliminate their traditional lectures and adopt an established method of individualizing instruction, even though they may prefer to teach with a traditional lecture? Since individualization implies a planning process it can be incorporated into any instructional system. It is possible for a traditional lecture to become more individualized with appropriate planning (e.g., the varying of assignments to meet different student needs, fexible testing, etc.). This may subsequently lead to a gain in the effectiveness of traditional lecture instruction because of "individualization."

Recommendations

The results of this meta-analysis tend to support efforts to implement individualized instruction in science. Additional research, however, should be encouraged in areas relating to the cost effectiveness. Does the small increase in effectiveness justify an increased expenditure for the implementation of an individualized instructional system?

The results of this meta-analysis lead to questions regarding the relationship between the effectiveness of individualized instruction and class size or teacher/student ratio. In this dissertation an attempt was made to identify teacher/student ratio as an instructional variable. This ratio was not, however, reported in most of the research studies analyzed. It would be a worthwhile endeavor to obtain this information by contacting the authors of the studies listed in Appendix A. This data may show that class size is the instructional variable which explains the variance in effect size demonstrated in this meta-analysis.

In addition to class size, other variables, such as study time and drop-out rate, should be examined in an attempt to explain the variance in the effect sizes reported in this dissertation. An investigation of the amount of time which students study may show that students in individualized courses study more than students in traditional courses, therefore explaining the difference in effect size. A study of drop-out rates may show that there is a higher drop-out rate in individualized courses than in traditional courses. This may lead to the selection of stronger students in individualized courses.

An examination of the Hawthorne effect should also be undertaken. The Hawthorne effect could be indexed by previous exposure to individualized courses. Perhaps it is the student response to the uniqueness of the individualized instructional process which results in the effect sizes obtained in the present study.

REFERENCES

Athappilly, K. A meta-analysis of the effects of modern mathematics in comparison with traditional mathematics. Paper presented at the meeting of the American Educational Research Association, Boston, April 1980.

Baez, V. Innovation in science education - world-wide. Paris: The Unesco Press, 1976.

Beauchamp, W.L. Instruction in science (Bulletin No. 17, National Survey of Secondary Education, Monograph No. 22). Washington, D.C.: U.S. Government Printing Office, 1932.

Belzer, T.J., & Conti, M.R. A comparison of grades achieved, attrition rates, and teacher evaluation in a traditional general biology course versus a non-traditional, audio-tutorial biology course. Nova University, 1973. (ERIC Document Reproduction Service No. ED 094 810)

Billett, R.O. Provisions for individual differences, marking and promotion (Bulletin No. 17, National Survey of Secondary Education, Monograph No. 13). Washington, D.C.: U.S. Government Printing Office, 1932.

Block, J.H., & Burns, R.B. Mastery learning. In L.S. Shulman (Ed.), Review of research in education (Vol. 4). Itasca, Ill.: Peacock, 1976.

Chambers, J.A., & Sprecher, J.W. Computer assisted instruction: Current trends and critical issues. Communication of the Association for Computing Machinery, 1980, 23, 332-342.

Cohen, J. Statistical power analysis for the behavioral sciences. New York: Academic Press, 1969.

Costin, F. Lecturing versus other methods of teaching: A review of research. British Journal of Educational Technology, 1972, 3(1), 4-31.

Dewey, J. Experience and education. New York: Macmillan Publishing Co., Inc., 1938.

Dubin, R., & Hedley, R.A. The medium may be related to the message: College instruction by TV. Eugene, Oregon: CASEA, 1969.

Dubin, R., & Taveggia, T.C. The teaching-learning paradox. Eugene Oregon: University of Oregon Press, 1968.

Eysenik, H.J. An exercise in mega-silliness. American Psychologist, 1978, 33, 517.

Fisher, K.M. A-T science teaching: How effective is it. Bioscience, 1976, 26, 691-697.

Fisher, K.M., & MacWhinney, B. AV autotutorial instruction: A review of evaluative research. AV Communication Review, 1976, 24, 229-261.

Gallo, P.S. Meta-analysis - a mixed meta-phor? American Psychologist, 1978, 33, 515-517.

Glass, G.V. Primary, secondary, and meta-analysis of research. Educational Researcher, 1976, 5(10), 3-8.

Glass, G.V. Integrating findings: The meta-analysis of research. In L.S. Shulman (Ed.) Review of research in education (Vol. 5), Itasca, Ill.: Peacock, 1978a.

Glass, G.V. Reply to Mansfield and Busse. Educational Researcher, 1978b, 7(1), 3.

Glass, G.V. Personal communication, 1979.

Glass, G.V, & Smith, M.L. Meta-analysis of research on the relationship of class size and achievement. Boulder, Colorado: Laboratory of Educational Research, University of Colorado, September, 1978.

Grobe, C.H. Comparisons in college biology achievement between the audio-tutorial and conventional methods of instruction of non-science majors (Doctoral dissertation, University of North Dakota, 1970). Dissertation Abstracts International, 1971, 32, 6824A. (University Microfilms No. 72-16361)

Hall, J.A. Gender effects in decoding non verbal cues. Psychological Bulletin, 1978, 85, 845-857.

Hartley, S.S. Meta-analysis of the effects of individually paced instruction in mathematics (Doctoral dissertation, University of Colorado, 1977). Dissertation Abstracts International, 1978, 38, 4003A. (University Microfilms No. 77-29926)

Hurd, P.D. Research in science education: Planning for the future. Journal of Research in Science Teaching, 1971, 8, 243-249.

Hursh, D.E. Personalized systems of instruction: What do the data indicate? Journal of Personalized Instruction, 1976, 1, 91-105.

Jamison, D., Suppes, P., & Wells, S. The effectiveness of alternative instructional media: A survey. Review of Educational Research, 1974, 44, 1-67.

Johnson, K.R., & Ruskin, R.S. Behavioral instruction: An evaluative review. Washington, D.C.: American Psychological Association, 1977.

Keller, F.S. Good-bye, teacher. Journal of Applied Behavior Analysis, 1968, 1, 79-88.

Kulik, J.A. PSI: A formative evaluation. In B.A. Green, Jr. (Ed.), Personalized instruction in higher education: Proceedings of the second national conference. Washington, D.C.: Center for Personalized Instruction, 1976.

Kulik, J.A., Cohen, P.A., & Ebeling, B.J. Effectiveness of programmed instruction in higher education: A meta-analysis of findings. Ann Arbor, MI: Center for Research on Learning and Teaching, The University of Michigan, September, 1979.

Kulik, J.A., Kulik, C.C., & Carmichael, K. The Keller plan in science teaching. Science, 1974, 183, 379-383.

Kulik, J.A., Kulik, C.C., & Cohen, P.A. A meta-analysis of outcome studies of Keller's personalized system of instruction. American Psychologist, 1979a, 34, 307-318.

Kulik, J.A., Kulik, C.C., Cohen, P.A. Effectiveness of computer-based college teaching: A meta-analysis of findings. Ann Arbor, MI: Center for Research on Learning and Teaching, The University of Michigan, 1979b.

Kulik, J.A., Kulik, C.C., & Cohen, P.A. Research on audio-tutorial instruction: A meta-analysis of comparative studies. Research in Higher Education, 1979c, 11, 321-341.

Kulik, J.A., Kulik, C.C., & Smith, B.B. Research on the personalized system of instruction. Programmed Learning and Educational Technology, 1976, 13, 23-30.

Lange, P.C. Today's education. National Education Association, 1972, 61, 59.

Light, R.J. Capitalizing on variation: How conflicting research findings can be helpful for policy. Educational Researcher, 1979, 8(9), 7-12.

Light, R.J., & Smith, P.V. Accumulating evidence: Procedures for resolving contradictions among different research studies. Harvard Educational Review, 1971, 41, 429-471/

Mallinson, G.G. A summary of research in science education - 1975. Washington, D.C.: National Institute of Education, 1976. (ERIC Document Reproduction Service No. ED 148 602)

Mansfield, R.S., & Busse, T.V. Meta-analysis of research: A rejoinder to Glass. Educational Researcher, 1977, 6(9), 3.

Marchese, R.S. A literature search and review of the comparison of individualized and conventional modes of instruction in science. School Science and Mathematics, 1977, 77, 699-703.

McGaw, B., & Glass, G.V. Choice of the metric for effect size in meta-analysis. American Educational Research Journal, 1980, 17, 325-337.

McKeachie, W.J., & Kulik, J.A. Effective college teaching. In F.N. Kerlinger (Ed.), Review of research in education, Vol. 3, Itasca, Ill.: F.E. Peacock, 1975, 165-209.

Meleca, C.B. Multiple linear regression analysis: Results and discussion II. Bioscience, 1970, 20, 26-30.

Mintzes, J.J. The A-T approach 14 years later - a review of recent research. Journal of College Science Teaching, 1975, 4, 247-252.

Nie, N.H., Hull, C.H., Jenkins, J.G., Steinbrenner, K., & Brent, D.H. Statistical package for the social sciences. New York: McGraw-Hill, 1975.

Pflaum, S.W., Walberg, H.J., Karegianes, M.L., & Rasher, S.P. Reading instruction: A quantitative analysis. Educational Researcher, 1980, 9(7), 12-18.

Postlethwait, S.N. A systems approach to biology. Audio-visual Instruction, 1963, 8, 243.

Razik, T.A. (Ed.) Bibliography of programmed instruction and computer assisted instruction. Englewood Cliffs, New Jersey: Educational Technology Publications, 1971.

Robin, A.R. Behavioral instruction in the college classroom. Review of Educational Research, 1976, 46, 313-354.

Rosenthal, R. Experimenter effects in behavioral research. New York: Irvington Publishers, 1976.

Rosenthal, R. Combining results of independent studies. Psychological Bulletin, 1978, 85, 185-193.

Russell, W.B. Some comparisons of the audio-tutorial method with the conventional method in introductory college biology (Doctoral dissertation, North Texas State University, 1968). Dissertation Abstracts International, 1968, 29, 3535A. (University Microfilms No. 69-05279)

Schramm, W. The research on programmed instruction: An annotated bibliography. Washington, D.C.: U.S. Department of Health, Education, and Welfare, 1964.

Seal, C.W. Two methods of teaching a general biology course in a community college: Audio-tutorial with conventional lecture-laboratory (Doctoral dissertation, Auburn University, 1976). Dissertation Abstracts International, 1976, 37, 5729A. (University Microfilms No. 77-04317)

Smith, M.L., & Glass, G.V. Meta-analysis of psychotherapy outcome studies. American Psychologist, 1977, 32, 752-760.

Smith, M.L., & Glass, G.V. Relationship of class-size to classroom processes, teacher satisfaction and pupil affect: A meta-analysis. Boulder, Colorado: Laboratory of Educational Research, University of Colorado, July, 1979.

Taveggia, T.C. Personalized instruction: A summary of comparative research, 1967-1974. American Journal of Physics, 1976, 44, 1028-1033.

√Taveggia, T.C. Goodbye teacher, goodbye classroom, hello learning: A radical appraisal of teaching-learning linkages at the college level. Journal of Personalized Instruction, 1977, 2, 119-122.

√ Triezenberg, H.J., & McLeod, R.J. Individuals learn. In H.J. Triezenberg, (Ed.), Individualized science. Like it is. Washington, D.C.: National Science Teachers Association, 1972.

Tyack, D.B. The one best system: A history of American urban education. Cambridge, Mass: Harvard University Press, 1974.

VanderWal, J.M. The relationships between two methods of teaching high school biology (Doctoral dissertation, Iowa State University, 1970). Dissertation Abstracts International, 1970, 31, 4637A. (University Microfilms No. 70-26,455)

Weaver, G.M. An evaluation and analysis of audio-tutorial instruction in a general education biology course (Doctoral dissertation, East Texas State University, 1969). Dissertation Abstracts International, 1969, 30, 5258A. (University Microfilms No. 70-11306)

Welch, W.W. Review of research 1968-69 in secondary level science. Journal of Research in Science Teaching, 1972, 9, 97-122.

APPENDIX A

STUDIES INCORPORATED IN META-ANALYSIS

Aaron, G. The effectiveness of programmed instruction when used to supplement or supplant assignments in biology classes in which team-teaching techniques are employed (Doctoral dissertation, University of Wisconsin-Madison, 1965). _Dissertation Abstracts International_, 1965, _25_, 6380. (University Microfilms No. 65-05108)

Anderson, C.J. The development and evaluation of programmed learning for high school chemistry (Doctoral dissertation, Arizona State University, 1971). _Dissertation Abstracts International_, 1971, _32_, 1935A. (University Microfilms No. 71-23007)

Anderson, O.T., & Artman, R.A. A self-paced, independent study, introductory physics sequence - description and evaluation. _American Journal of Physics_, 1972, _40_, 1737-1742.

Arlin, M.H. The leveling effect of teacher pacing on science content mastery. _Journal of Research in Science Teaching_, 1976, _13_, 213-219.

Austin, S.N., & Gilbert, K.E. Student performance in a Keller plan course in introductory electricity and magnetism. _American Journal of Physics_, 1973, _41_, 12-18.

Banks, J.E. Evaluation of an auto-instructional program in general college chemistry. _Journal of Research in Science Teaching_, 1963, _1_, 351-352.

Becker, W.A., & Shumway, L.K. Innovative methods of learning in a general genetics course. _Journal of Heredity_, 1972, _63_, 122-128.

Belzer, T.J., & Conti, M.R. _A comparison of grades achieved, attrition rates, and teacher evaluation in a traditional general biology course versus a non-traditional, audio-tutorial biology course_. Nova University, 1973. (ERIC Document Reproduction Service No. ED 094 810)

Boren, A.R., & Foree, S.B. Personalized instruction applied to food and nutrition in higher education. Journal of Personalized Instruction, 1977, 2, 39-42.

Brantley, W.T. A comparison of the audio-tutorial method with the lecture-demonstration method for producing student achievement in college level physical science and astronomy (Doctoral dissertation, University of Southern Mississippi, 1974). Dissertation Abstracts International, 1974, 35, 6526A. (University Microfilms No. 75-09576)

Brewer, I.M. Recall, comprehension and problem solving evaluation of an audio-visual method of learning plant anatomy. Journal of Biological Education, 1974, 8, 101-112.

Brown, W.A. An experiment using programmed material as supplemental instruction in college physical science (Doctoral dissertation, University of Florida, 1963). Dissertation Abstracts, 1964, 25, 1041. (University Microfilms No. 64-5636)

Call, R.L. A comparison of individualized and traditional methods for teaching high school chemistry (Doctoral dissertation, Arizona State University, 1974). Dissertation Abstracts International, 1975, 35, 7140A. (University Microfilms No. 75-10038)

Carlson, J.A. The design and evaluation of an individualized, contract-directed high school chemistry course (Doctoral dissertation, Northern Illinois University, 1975). Dissertation Abstracts International, 1976, 36, 6580A. (University Microfilms No. 76-8906)

Castleberry, S.J., Montague, E.J., & Lagowski, J.J. Computer-based teaching techniques in general chemistry. Journal of Research in Science Teaching,1970, 7, 197-208.

Castleberry, S.J. The development and evaluation of computer-assisted instruction programs on selected topics in introductory college chemistry (Doctoral dissertation, University of Texas at Austin, 1970). Dissertation Abstracts International, 1970, 31, 1679A. (University Microfilms No. 70-18214)

Cheek, W.E. Comparison of traditional instruction in chemistry. Nova University, 1976. (ERIC Document Reproduction Service No. ED 131 871)

Combs, R.E. Student achievement through individualization. American Biology Teacher, 1975, 37, 171-173.

Cote, J.D. Biology by PSI in a community college. Proceedings of the Second National Conference held by the Center for Personalized Instruction, 1975, 67-72.

Cowan, P.J. Development of new autoinstructional materials and an analysis of their effectiveness in teaching modern physics in the small high school (Doctoral dissertation, University of Texas, 1964). Dissertation Abstracts, 1964, 25, 2879-2880. (University Microfilms No. 64-11787)

Cross, J.Z., & Semb, G. An analysis of the effects of personalized instruction on students at different initial performance levels in an introductory college nutrition course. Journal of Personalized Instruction, 1976, 1, 47-50.

Culp, G.H. An approach to the use of computer-assisted instruction in organic chemistry (Doctoral dissertation, University of Texas at Austin, 1969). Dissertation Abstracts International, 1969, 30, 5308A. (University Microfilms No. 70-10771)

Culp, G.H., Stotter, P.L., & Gilbert, J.C. Adapting computer-based instruction in undergraduate organic chemistry: An experimental course. University of Texas, 1973. (ERIC Document Reproduction Service No. ED 116 605)

Darnowski, V.S. Three types of programmed learning and the conventional teaching of the nuclear chemistry portion of the high school chemistry course (Doctoral dissertation, New York University, 1968). Dissertation Abstracts International, 1968, 29, 509A. (University Microfilms No. 68-11785)

DeBoer, G.E. A comparison of two methods of teaching chemistry to freshmen student nurses: An individualized approach versus lecture (Doctoral dissertation, Northwestern University, 1972). Dissertation Abstracts International, 1972, 33, 2610A. (University Microfilms No. 72-32419)

Denton, J.J. A methodological study of a computer-managed instructional program in high school physics (Doctoral dissertation, University of Missouri-Columbia, 1972). Dissertation Abstracts International,1972, 33, 4966. (University Microfilms No. 73-7023)

Durst, H.E. An evaluation of an experimental biology course employing some cognitive and non-cognitive achievement as criteria (Doctoral dissertation, Oregon State University, 1967). Dissertation Abstracts, 1968, 28, 2438A. (University Microfilms No. 67-17880)

Elliott, G.E. A study to determine if selected student characteristics are related to success in introductory physics using aptitude-treatment interaction analysis for two diverse modes of instruction (Doctoral dissertation, Univeristy of Maryland, 1976). Dissertation Abstracts International, 1977, 37, 7055A. (University Microfilms No. 77-9505)

Fisher, K.M., Guenther, H., & MacWhinney, B. Does video-autotutorial instruction improve college student achievement? Journal of Research in Science Teaching, 1977, 14, 481-498.

Fivaz, R. Physics for engineers: A European attempt at PSI. Journal of Personalized Instruction, 1977, 2, 156-161.

Francis, E.E. An analysis of achievement in a college biology audio-tutorial program (Doctoral dissertation, Arizona State University, 1976). Dissertation Abstracts International, 1976, 37, 5727A. (University Microfilms No. 77-05570)

Francis, G.H. An experimental study of the effectiveness of self-instruction versus the lecture-demonstration method of teaching selected phases of electricity (Doctoral dissertation, Southern Illinois University, 1966). Dissertation Abstracts, 1966, 27, 3338A. (University Microfilms No. 67-3157)

Fulton, H.E. An analysis of student outcomes utilizing two approaches to teaching BSCS biology (Doctoral dissertation, University of Iowa, 1970). Dissertation Abstracts International, 1970, 31, 1137A. (University Microfilms No. 70-15,600)

Fulton, H.E. Individualized vs. group teaching of BSCS biology. American Biology Teacher, 1971, 33, 277-279.

Fulton, H.F. A comparative study of student attitudes toward science and the ability of the teacher to make material understandable in individualized and group approaches to BSCS biology. School Science and Mathematics, 1971, 71, 198-202.

Garren, D.R.,& Gathmann, D.A. Audio-tutorial instruction and student selected curricula. Paper presented at National Conference on Behavior Research and Technology in Higher Education, Georgia State University, November, 1974. (ERIC Document Reproduction Service No. ED 097 890)

Geiger, H.B. Effectiveness in learning Newton's second law of motion in secondary school physics using three methods of learning (Doctoral dissertation, Temple University, 1968). Dissertation Abstracts International, 1968, 29, 1821. (University Microfilms No. 68-17823)

Geller, M. The measurement of the effectiveness of a teaching machine program in the organic chemistry area of first year college chemistry. Journal of Research in Science Teaching, 1963, 1, 154-161.

Gerrell, G.E. Computer assisted instruction of college physics students in small groups. Final report National Center for Educational Research and Development, Washington, D.C., 1972. (ERIC Document Reproduction Service No. ED 064 952)

Gould, J.C., Langfor, N.G., & Mott, C.J. Earth science as an audio-tutorial course. Journal of Geological Education, 1972, 20, 81-83.

Grandey, R.C. An investigation of the use of computer-aided instruction in teaching students how to solve selected multi-step general chemistry problems (Doctoral dissertation, University of Illinois at Urbana-Champaign, 1970). Dissertation Abstracts International, 1971, 31, 6430. (University Microfilms No. 71-14764)

Griffiths, A.K., & Crocker, R.K. Achievement in individualized and conventional chemistry courses and its interaction with selected academic and personality variables. The Alberta Journal of Educational Research, 1976, 12, 97-105.

Grobe, C.H. Comparisons in college biology achievement between the audio-tutorial and conventional methods of instruction of non-science majors (Doctoral dissertation, University of North Dakota, 1970). Dissertation Abstracts International, 1971, 32, 6824A. (University Microfilms No. 72-16361)

Hahn, T.C. Audio-tutorial instruction: A case study. Bioscience, 1971, 21, 814-819.

Hammach, J.K. A comparative study of the effects of self-paced and lecture laboratory instruction on achievement in general studies biology (Doctoral dissertation, East Texas State University, 1974). Dissertation Abstracts International, 1975, 35, 4187A. (University Microfilms No. 75-1582)

Hardin, L.D. A study of the influence of a physics personalized system of instruction versus lecture on cognitive reasoning, achievement, attitudes and critical thinking (Doctoral dissertation, University of Northern Colorado, 1977). Dissertation Abstracts International, 1977, 38, 4711A. (University Microfilms No. 77-30826)

Hedges, L. Personalized introductory courses: A longitudinal study. American Journal of Physics, 1978, 46, 207-210.

Heffley, P.D. The implementation of the personalized system of instruction in the freshman chemistry course at Canisius College. Proceedings of the First National Conference held by the Center for Personalized Instruction, 1974, 140-145.

Hug, W.E. An experiment comparing cognitive and affective dimensions of independent study, small-group instruction in high school biology (Doctoral dissertation, University of Southern California, 1969). Dissertation Abstracts International, 1969, 30, 618A. (University Microfilms No. 69-13,063)

Hunter, W.E. An analysis of a systems approach to the teaching-learning process (Doctoral dissertation, University of Colorado, 1972). Dissertation Abstracts International, 1973, 33, 3955. (University Microfilms No. 73-1786)

Hunter, W.E. Individualized approaches to chemistry vs. group lecture discussions. Journal of College Science Teaching, 1973, 2(4), 35-38.

Ijaz, L.K. Individualizing science education via the triple-stage interaction model (Doctoral dissertation, Virginia Polytechnic Institute and State University, 1975). Dissertation Abstracts International, 1975, 36, 2593. (University Microfilms No. 75-23702)

Inskeep, W.K. The effectiveness of a multiple media approach in teaching certain concepts in high school chemistry (Doctoral dissertation, Indiana University, 1969). Dissertation Abstracts International, 1969, 30, 619A. (University Microfilms No. 69-12, 524)

Jacobs, L.C. The influence of teaching machine procedures upon learning in high school chemistry (Doctoral dissertation, Indiana University, 1961). Dissertation Abstracts , 1961, 22, 3522. (University Microfilms No. 61-06942)

Kahle, J.B., Nordland, F.H., & Douglass, C.B. An analysis of an alternative instructional model for disadvantaged students. Science Education, 1976, 60(2), 237-243.

Kantasewi, N. Experiments in the use of programed materials in teaching an introductory course in the biological sciences at the college level (Doctoral dissertation, Pennsylvania State University, 1964). Dissertation Abstracts, 1964, 25, 2378-2379. (University Microfilms No. 64-10785)

Krockover, G.H. A comparison of learning outcomes in CBA chemistry when group and individualized instruction techniques are employed (Doctoral dissertation, University of Iowa, 1970). Dissertation Abstracts International, 1970, 31, 1240A. (University Microfilms No. 70-15614)

Kromhout, O.M. Effect of computer tutorial review lessons on exam performance in introductory college physics. Florida State University, 1972. (ERIC Document Reproduction Service No. ED 072 654)

Lagowski, J.J., Castleberry, S.J., & Culp, G.H. The impact of computer-based instructional methods in general chemistry. Proceedings of the 1972 Conference in Computers in Undergraduate Curricula, Georgia Institute of Technology, Atlanta, 1972, 77-99.

Lee, E.W. A study of the effect of two methods of teaching high school chemistry upon critical thinking abilities (Doctoral dissertation, University of North Carolina at Chapel Hill, 1963). Dissertation Abstracts International, 1964, 25, 4578. (Univerity Microfilms No. 64-9423)

Lewis, D.K., & Wolf, W.A. Implementation of self-paced learning (Keller plan) in a first year course. Journal of Chemical Education, 1973, 50, 51-56.

Lindsay, C.W. A study of the effects of three methods of teaching high school chemistry upon achievement in chemistry, critical thinking abilities, and scientific interest (Docotral dissertation, Northeast Louisiana University, 1973). Dissertation Abstracts International, 1974, 34, 7068A. (University Microfilms No. 74-11454)

Little, E.B., & Sackheim, G.I. A chemistry program for beginning biology. School Science and Mathematics, 1965, 65, 384-388.

Marinos, N.G., & Lucas, A.M. An assessment of audio-tutorial teaching of a plant physiology course. Journal of Biological Education, 1971, 5, 109-113.

Marshall, G. The development and evaluation of a programed supplementary guide for selected topics in high school biology (Doctoral dissertation, University of Florida, 1970). Dissertation Abstracts International, 1970, 31, 5875A. (University Microfilms No. 71-12,763)

McKee, R.J. A comparative study of two programed instructional methods and conventional instruction in a unit of ninth grade physical science (Doctoral dissertation, University of North Dakota, 1966). Dissertation Abstracts, 1966, 27, 3371A. (University Microfilms No. 67-4466)

Meleca, C.B. Multiple linear regression analysis: Results and discussion II. Bioscience, 1970, 20, 26-30.

Miller, W.R. Identification of factors related to student achievement and attitude in two methods of instruction for first semester college chemistry (Doctoral dissertation, University of Texas at Austin, 1974). Dissertation Abstracts International, 1974, 35, 2794A. (University Microfilms No. 74-24906)

Mintzes, J.J. A comparative study of student characteristics as correlates of performance in a college biology program employing audio-tutorial teaching modules (Doctoral dissertation, Northwestern University, 1974). Dissertation Abstracts International, 1974, 35, 3543A. (University Microfilms No. 74-28689)

Moriber, G. The measurement of the effectiveness of programmed learning in an area of chemistry in a physical science course for non-science college majors (Doctoral dissertation, New York University, 1967). Dissertation Abstracts International, 1967, 28, 1348A. (University Microfilms No. 67-11117)

Moore, J.W., Hauck, W.E., & Gagne, E.D. Acquisition, retention, and transfer in an individualized college physics course. Journal of Educational Psychology, 1973, 64, 335-340.

Moore, J.W., Mahan, J.H., & Ritt, C.A. Continuous progress concept of instruction with university students. Psychological Reports, 1969, 25, 887-892.

Oines, R.K. The comparative effectiveness of individually prescribed instruction and the lecture demonstration method to achieve behavioral objectives for a descriptive astronomy course (Doctoral dissertation, Oklahoma State University, 1971). Dissertation Abstracts International, 1971, 33, 666. (University Microfilms No. 72-21960)

Ott, M.D., & Macklin, D.B. A trait-treatment interaction in a college physics course. Journal of Research in Science Teaching, 1975, 12, 111-119.

Paden, J.S. An experimental study of individualized instruction in high school physics using the computer to prescribe activities as a function of selected idiographic factors (Doctoral dissertation, University of Missouri-Columbia, 1970). Dissertation Abstracts International, 1970, 31, 5275A. (University Microfilms No. 71-08373)

Parker, B.D. An investigation of aptitude-treatment interaction using two distinct college chemistry instructional modes and a selected set of learner aptitudes (Doctoral dissertation, Syracuse University, 1973). Dissertation Abstracts International, 1974, 35, 1521A. (University Microfilms No. 74-10169)

Parker, G.E. The relationship of programmed instruction to test and discussion performance among beginning college biology students (Doctoral dissertation, Ball State University, 1973). Dissertation Abstracts International, 1973, 34, 4914A. (University Microfilms No. 74-1726)

Philippas, M.A., & Sommerfeldt, R.W. Keller vs. lecture method in general physics instruction. American Journal of Physics, 1972, 40, 1300-1306.

Protopapas, P. A report on the use of the Keller plan in a general biology course at Lowell State College. In J.G. Sherman (Ed.), PSI Personalized System of Instruction, 41 Germinal Papers, Mendo Park, California: W.A. Benjamin, Inc., 1974.

Redditt, R.S. A quasi-experimental comparison of a group lecture method and a self-directed method in teaching basic electricity at the college level (Doctoral dissertation, University of Tennessee, 1973). Dissertation Abstracts International, 1974, 34, 5599A. (University Microfilms No. 74-3861)

Reed, L.H. The effects of individualized instruction in science upon the achievement, attitude, and self-concept of inner-city secondary students (Doctoral dissertation, North Texas State University, 1974). Dissertation Abstracts International, 1975, 35, 7748A. (University Microfilm No. 75-13681)

Richard, P., & Sund, R.B. Individualized instruction in biology. American Biology Teacher, 1969, 31, 252-256.

Ricker, P.E. An experimental comparison of four methods of presenting basic properties of magnetism (Doctoral dissertation, Colorado State College, 1965). Dissertation Abstracts, 1965, 26, 5306. (University Microfilms No. 65-14,830)

Ritter, W.R. A comparison of achievement for two methods of instruction with the use of behavioral objectives (Doctoral dissertation, University of Pennsylvania, 1975). Dissertation Abstracts International, 1975, 36, 4384A. (University Microfilms No. 75-26130)

Rodewald, L.B., Culp, G.H., & Lagowski, J.J. The use of computers in organic chemistry instruction. Journal of Chemical Education, 1970, 47, 134-136.

Rosati, P.A. A comparison of the personalized system of instruction with the lecture method in teaching elementary dynamics. In J.M. Johnson Behavior research and technology in higher education. Springfield, Illinois: Charles C. Thomas, Publisher, 1975.

Rowsey, R.E., & Mason, W.H. Immediate achievement and retention in audio-tutorial versus conventional lecture-laboratory instruction. Journal of Research in Science Teaching, 1975, 12, 393-397.

Ruiter, W.W. An analysis of the effectiveness of a learner-centered teaching system compared to that of a conventional teaching of basic electricity to university students (Doctoral dissertation, Oregon State University, 1971). Dissertation Abstracts International, 1971, 32, 1926A. (University Microfilm No. 71-25072)

Russell, W.B. Some comparisons of the audio-tutorial method with the conventional method in introductory college biology (Doctoral dissertation, North Texas State University, 1968). Dissertation Abstracts International, 1968, 29, 3535A. (University Microfilms No. 69-05279)

Sasscer, J.C. The development, implementation, and evaluation of a modularized, student-centered general biology curriculum at the college level (Doctoral dissertation, Catholic University, 1973). Dissertation Abstracts International, 1974, 34, 6957. (University Microfilms No. 74-11452)

Seal, C.W. Two methods of teaching a general biology course in a community college: Audio-tutorial with conventional lecture-laboratory (Doctoral dissertation, Auburn University, 1976). Dissertation Abstracts International, 1976, 37, 5729A. (University Microfilms No. 77-04317)

Siemankowski, F.T. An auto-paced teaching process in physical science for elementary teacher preparation: A pilot report. Journal of Research in Science Teaching, 1969, 6, 150-156.

Silberman, R., & Parker, B. Student attitudes and the Keller plan. Journal of Chemical Education, 1974, 51, 393-395.

Smith, C.J. A comparative study of two methods of teaching physical science to college freshmen with dis-advantaged backgrounds (Doctoral dissertation, University of Mississippi, 1970). Dissertation Abstracts International, 1970, 31, 1152A. (University Microfilms No. 70-16406)

Sparks, P.D., & Unbehaum, L.M. Achievement of audio-tutorial and conventional biology students - a comparative study. Bioscience, 1971, 21, 574-576.

Spevack, H.M. A comparison of the personalized system of instruction with the lecture recitation system for nonscience oriented chemistry students at an open enrollment community college (Doctoral dissertation, New York University, 1975). Dissertation Abstracts International, 1975, 36, 4385A. (University Microfilms No. 76-01757)

Staniskis, C.C. A comparison of student content achievement in biology between computer managed instructional and non computer managed instructional biology courses (Doctoral dissertation, Temple University, 1976). Dissertation Abstracts International, 1977, 37, 7665A. (University Microfilms No. 77-13528)

Steele, R.W. Development and evaluation of an individualized physics course for non science majors (Doctoral dissertation, University of Wyoming, 1974). Dissertation Abstracts International, 1975, 35, 4274A (University Microfilms No. 75-238)

Strickland, W.R. A comparison of a programed course and a traditional lecture course in general biology (Doctoral dissertation, University of Southern Mississippi, 1971). Dissertation Abstracts International, 1971, 32, 2510A. (University Microfilms No. 71-28849)

Summerlin, L., & Gardner, M. A study of tutorial type computer assisted instruction in high school chemistry. Journal of Research in Science Teaching, 1973, 10(1), 75-82.

Tamminen, M. The effect of a programmed supplement of general chemistry problems on the problem solving skills of college chemistry students (Doctoral dissertation, Boston College, 1976). Dissertation Abstracts International, 1976, 36, 7320A. (University Microfilms No. 76-10393)

Thorsland, M.N. Formative evaluation in the development of an audio-tutorial physics course. Science Education, 1975, 59, 305-312.

Tookey, J.V. Evaluating the effectiveness of a linear programed course in science education. Journal of Research in Science Teaching, 1964, 2, 59.

Torop, W. An analysis of individualized learning system chemistry. Paper presented at annual meeting of National Association for Research in Science Teaching, Los Angelos, March, 1975. (ERIC Document Reproduction Service No. ED 110 321)

VanderWal, J.M. The relationships between two methods of teaching college biology in achievement and attitude (Doctoral dissertation, Western Michigan University, 1972). Dissertation Abstracts International, 1973, 34, 215A. (University Microfilms No. 73-15038)

Volker, R.P. Development of a multimedia system for teaching high school biology (Doctoral dissertation, Iowa State University, 1970). Dissertation Abstracts International, 1970, 31, 4051A. (University Microfilms No. 70-25833)

Waine, S.I. The effectiveness of a programmed textbook in teaching selected chemistry topics to high school introductory biology students (Doctoral dissertation, New York University, 1970). Dissertation Abstracts International, 1970, 31, 4637A. (University Microfilms No. 70-26,455)

Walsh, R.G. The Keller plan in college introductory physical geology: A comparison with the conventional teaching method (Doctoral dissertation, University of North Dakota, 1975). Dissertation Abstracts International, 1976, 37, 4257A. (University Microfilms No. 76-3029)

Wash, J.A. An experiment in the use of programed materials in teaching high school chemistry (with) rates of reaction and chemical equilibrium: A programed sequence (Doctoral dissertation, University of Georgia, 1964). Dissertation Abstracts, 1965, 26, 914-915. (University Microfilms No. 65-4509)

Weaver, G.M. An evaluation and analysis of audio-tutorial instruction in a general education biology course (Doctoral dissertation, East Texas State University, 1969). Dissertation Abstracts International, 1969, 30, 5258A. (University Microfilms No. 70-11306)

Weiss, E. An investigation of the relative effectiveness of two techniques of teaching a transitional course in physics on the college level (Doctoral dissertation, New York University, 1971). Dissertation Abstracts International, 1971, 32, 2411A. (University Microfilms No. 71-28569)

Welker, J.D. Selected factors and achievement in an audio-tutorial introductory college biology course (Doctoral dissertation, Indiana University, 1969). Dissertation Abstracts International, 1970, 31, 249A. (University Microfilms No. 70-11710)

Wheaton, W.D. An evaluation of an individualized learning program in a California University high school district (Doctoral dissertation, University of Southern California, 1971). Dissertation Abstracts International, 1972, 32, 5540A. (University Microfilms No. 72-11966)

White, J.M., Close, J.S., & McAllister, J.W. Freshman chemistry without lectures. Journal of Chemical Education, 1972, 49, 772-774.

Williams, W.W. An experimental investigation of individualized instruction in the teaching of quantitative physical science (Doctoral dissertation, Duke University, 1969). Dissertation Abstracts International, 1970, 31, 312A. (University Microfilms No. 70-8839)

Whilton, P.J., Whittle, S.J., & Teather, D.C.B. A comparison between teaching methods in secondary school biology. Journal of Biological Education, 1971, 5, 179-185.

Wright, J.C. Evaluation of an individualized earth science course designed for college non-science majors (Doctoral dissertation, University of Iowa, 1972). Dissertation Abstracts International, 1973, 33, 6750A. (University Microfilms No. 73-13606)

Young, A.F. A comparative study of supplementary programmed and conventional methods of instruction in teaching freshman chemistry 1015 at Oklahoma State University (Doctoral dissertation, Oklahoma State University, 1970). Dissertation Abstracts International, 1970, 31, 5882A. (University Microfilms No. 71-11311)

Young, P.A. An experiment in the use of programmed materials in teaching high school biology (Doctoral dissertation, Georgia University, 1967). Dissertation Abstracts International, 1967, 28, 4552. (University Microfilms No. 67-16246).

APPENDIX B

CORRELATION COEFFICIENTS

Table A

Correlation Coefficients of Method of Instruction for All Outcomes

ES	A-T	CAI	PSI	PI	Variable	S.D.
1.000	-.137	.033	.050	-.088	ES	.556
	1.000	-.146	-.224	-.301	A-T	.412
		1.000	-.119	-.160	CAI	.259
			1.000	-.246	PSI	.363
				1.000	PI	.433

Table B

Correlation Coefficients of Method of Instruction for Achievement

ES	A-T	CAI	PSI	PI	Variable	S.D.
1.000	-.157	.120	-.008	.009	ES	.558
	1.000	-.184	-.236	-.306	A-T	.428
		1.000	-.138	-.180	CAI	.300
			1.000	-.230	PSI	.360
				1.000	PI	.423

Table C

Correlation Coefficients of Instructional Variables for All Outcomes

ES	Mastery	Self-test	No choice of Deliv. Syst.	Pace= Daily	Pace= Weekly	Variable	S.D.
1.000	.205	-.116	-.136	-.042	-.039	ES	.534
	1.000	-.606	.009	-.202	-.367	Mastery	.439
		1.000	.147	.275	.410	Self-test	.483
			1.000	.215	-.054	No choice	.413
				1.000	-.370	Daily	.352
					1.000	Weekly	.500

Table D

Correlation Coefficients of Instructional Variables for Achievement

ES	Mastery	Self-test	No choice of Deliv. Syst.	Pace= Daily	Pace= Weekly	Variable	S.D.
1.000	.177	.000	-.119	.031	-.068	ES	.525
	1.000	-.560	-.013	-.174	-.356	Mastery	.435
		1.000	.197	.241	.430	Self-test	.478
			1.000	.219	.075	No choice	.417
				1.000	-.380	Daily	.353
					1.000	Weekly	.501

Table E

Correlation Coefficients of Design Variables for All Outcomes

ES	Self-Select	Instr. Cont.	Hist. Effect	Rating Exc.	Rating Good	Rating Fair	Inst. Teach.	Devel. Team	Equivalence of Subjects Absent	1 meas.	2+ meas.	Cov.	Variable	S.D.
1.000	-.236	-.001	.036	.168	-.019	-.041	.060	.067	.152	-.255	-.023	.119	ES	.565
	1.000	-.149	.158	.108	.030	-.105	-.109	.003	-.289	-.066	.129	-.089	Self-sel	.341
		1.000	-.193	-.100	.103	-.187	-.047	.169	.110	.152	.105	-.171	Ins cont	.502
			1.000	.416	-.094	.245	-.186	-.072	-.008	-.006	-.022	.354	Hist.eff	.349
				1.000	-.220	-.199	-.218	-.049	-.100	-.125	-.132	.469	Exc.	.257
					1.000	-.579	-.141	.042	-.242	-.195	.026	.113	Good	.490
						1.000	.054	.090	-.008	.106	.074	-.202	Fair	.479
							1.000	-.475	.213	.073	-.128	.086	Teach.	.502
								1.000	-.107	.110	.088	-.034	Team	.385
									1.000	-.166	-.175	-.184	Abs.	.323
										1.000	-.219	-.230	1 meas	.379
											1.000	-.243	2+ meas.	.392
												1.000	Cov	.404

Table F

Correlation Coefficients of Design Variables for Achievement

	Self-	Instr.	Hist.	Rating			Inst. Devel.		Equivalence of Subjects					
ES	Select	Cont.	Effect	Exc.	Good	Fair	Teach.	Team	Absent	1 meas.	2+ meas.	Cov.	Variable	S.D.
1.000	-.317	.080	-.046	-.091	-.003	-.046	.091	.090	.219	-.168	-.057	.035	ES	.529
	1.000	-.165	.175	.096	.075	-.105	-.124	.052	-.296	-.107	.114	-.017	Self-sel	.364
		1.000	-.136	0	.073	-.175	-.096	.148	.136	.136	.093	-.174	Ins cont	.503
			1.000	.228	-.050	-.225	-.107	-.036	.028	.028	.076	.201	Hist.eff	.352
				1.000	-.180	-.162	-.246	.026	-.091	-.091	-.104	.428	Exc.	.214
					1.000	-.584	-.199	.019	-.259	-.119	.007	.115	Good	.491
						1.000	.107	.070	-.082	.061	.119	-.135	Fair	.478
							1.000	-.554	.234	.098	-.138	-.108	Teach.	.501
								1.000	-.121	.048	.075	.026	Team	.404
									1.000	-.167	-.190	-.213	Abs.	.352
										1.000	-.190	-.213	1 meas	.352
											1.000	-.243	2+ meas.	.385
												1.000	Cov	.413

Table G

Correlation Coefficients of Miscellaneous Variables for All Outcomes

			Source			Setting		Subject				
ES	Level	Repl	Journ	Doc	Diss	Sec.Sh	Comm.	Biol	Chem	Phys	Variable	S.D.
1.000	.132	.021	.023	.063	-.061	-.133	.032	-.061	.009	.058	ES	.556
	1.000	.027	.185	-.032	-.208	-.231	-.121	.107	-.043	.049	Level	.308
		1.000	-.100	.213	-.007	-.116	-.107	-.168	.093	-.004	Repl	.277
			1.000	-.183	-.788	-.183	-.183	.060	-.117	.107	Journ	.452
				1.000	-.364	-.195	.029	-.070	.155	.006	Doc	.269
					1.000	.315	.137	-.026	.038	-.130	Diss	.489
						1.000	-.238	-.009	.230	-.164	Sec Sch	.464
							1.000	.205	-.055	-.092	Comm	.315
								1.000	-.481	-.355	Biol	.471
									1.000	-.351	Chem	.469
										1.000	Phys	.405

Table H

Correlation Coefficients of Miscellaneous Variables for Achievement

			Source			Setting			Subject			
ES	Level	Repl	Journ	Doc	Diss	Sec.Sh	Comm.	Biol	Chem	Phys	Variable	S.D.
1.000	.139	.116	.013	.085	-.075	-.113	.005	-.092	.081	.037	ES	.559
	1.000	.044	.161	-.012	-.171	-.244	-.131	.113	-.029	.045	Level	.332
		1.000	-.103	.216	.002	-.097	-.120	-.198	.194	-.003	Repl	.311
			1.000	-.195	-.776	-.171	-.166	.075	-.080	.012	Journ	.462
				1.000	-.348	-.191	.004	-.083	.148	.041	Doc	.273
					1.000	.311	.119	-.046	.004	-.050	Diss	.496
						1.000	-.224	-.032	.142	-.083	Sec Sch	.458
							1.000	.164	-.053	-.080	Comm	.311
								1.000	-.513	-.337	Biol	.481
									1.000	-.311	Chem	.469
										1.000	Phys	.377

The two page vita has been removed from the scanned document. Page 1 of 2

The two page vita has been removed from the scanned document. Page 2 of 2

A META-ANALYSIS COMPARING ALTERNATIVE METHODS OF INDIVIDUALIZED AND TRADITIONAL INSTRUCTION IN SCIENCE

by

Nancy Carol Aiello

(ABSTRACT)

During the past fifty years many educational researchers have investigated the effectiveness of alternative methods of instruction in science. These investigations have resulted in the publication of numerous studies which frequently report contradictory findings. The purpose of this study was to synthesize the studies conducted on individualized instruction in secondary school and collegiate science courses through a meta-analysis, the statistical integration and analysis of research studies. The major goals of the meta-analysis were to determine the experimental effect of individualized instruction in science and to explain the variation in the effect sizes obtained.

This dissertation included an analysis of 115 studies comparing individualized instruction with traditional lecture instruction. The methods of individualized instruction incorporated were audio-tutorial instruction, computer-assisted instruction, Keller's personalized

system of instruction and programmed instruction. Instruction which was individualized but not clearly identified as one of the previously listed methods was included under a combination category. The effectiveness of individualized instruction was measured in terms of effect size, a standardized difference in means between the individualized and the lecture classes. Effect sizes were calculated using the pooled standard deviation.

A number of instructional, design and miscellaneous variables were identified in the research studies. Instructional variables identified were mastery orientation, self-initiated testing, self-pacing of instruction, and choice of instructional delivery systems. Design variables identified were self-selection, historical effect, subjective rating of study, instrument development, equivalence of subjects, and instrument reliability. Miscellaneous variables identifed were level of instruction, nature of instruction, source of publication, year of publication, setting of the study, and subject of instruction.

Data analysis involved: (a) descriptive statistics for the entire data set, (b) descriptive statistics for each of the individualized instructional methods studied, (c) analysis of variance and t-tests to test the effect of selected variables on instructional effectiveness, and (d) regression analysis in which effect sizes were regressed onto selected independent variables.

The findings of this meta-analysis suggest that individualized instruction in science is somewhat more effective than traditional instruction. All five methods of individualized instruction analyzed were not differentially effective when compared to traditional instruction. The instructional, design, and miscellaneous variables identified in this meta-analysis contributed little to the variance in effect size.