CHAPTER 1: PURPOSES, PROCESSES, AND DOMAINS OF EDUCATIONAL RESEARCH

Terms presented and discussed in Chapter 1

quantitative research	qualitative research	document analysis research	program and policy research
basic research	applied research	action research	experimental research
correlation	regression	analysis of covariance	factor analysis
causal-comparative study	ex post facto study	developmental study	longitudinal study
cross-sectional study	a priori observations	naturalistic research	qualitative research
the *field*	participant observation	non-participant observation	case study
research criticism	critical research	evaluation	assessment

Overview of Chapter 1

Chapter 1 begins with a statement of the rationale and purpose of the book. The first section argues why critical reading of statistical research is important—for both potential professors and researchers and for those in teaching or administrative leadership positions in public or private schools. After demonstrating the predominance of statistical research within the total realm of educational research, the chapter then presents a statement of the rationale used in this book for teaching statistical knowledge. That rationale is, basically, that learning to read statistical research is both a verbal learning task and a visual or spatial learning task, but definitely is not a mathematical learning task. The following section then defines, loosely, the parameters of educational research and includes a lengthy description of the organization of the varieties of educational research. The chapter describes how there are more domains of educational research than statistical and though this book does not cover those topics, students must build into their schema of the organization of educational research those non-statistical research genres. The next section describes the general organizational units of the research report (which will be greatly expanded in Chapter 7) and the general standards of high-quality research, most specifically, statistical research. The final section is a refrain, again demonstrating the importance of becoming a critical reader of educational research.

THE CRITICAL ANALYSIS OF RESEARCH

Regardless of career goals, the first competencies that should be learned by those engaged in the study of educational research are learning to comprehend, analyze, summarize, evaluate, and critique the written research found in the full spectrum of research documents available in education: journals, monographs, books, and unpublished manuscripts. Such competencies are not come upon naturally however, and their development requires instruction, guidance, and practice. Educational research does not come in just one variety either, indeed, educational research can be classified into four domains: quantitative, qualitative, document analysis, and program and policy

research (later sections of this chapter will specify these domains). To some degree, varying knowledge and abilities are needed to be a critical reader in these different domains; however, no educator seeking to be an educational scholar will be able to evade the reading of statistical research—as statistical research predominates the literature. Elmore & Woehlke (1988) classified the 549 articles that appeared in the *American Educational Research Journal* between 1978 and 1987 into one of 14 categories; 10 categories and 502—91%—of the studies were quantitative. Similarly, Baldwin et al. (1992) analyzed 1,848 research articles (plus 291 essays) that appeared in the *Journal of Reading Behavior* between 1952 and 1991, and 1,668—90%—were quantitative. To become an educational scholar, one must be able to comprehend and evaluate statistical research—and this is the purpose of this book.

Traditionally, doctoral programs have introduced students to statistical research by requiring them to enroll in educational statistics courses—where the focus is on learning statistics in order to do statistical analyses—not on the critical analysis of already completed statistical research. Learning how to do statistics does not directly apply to the critical analysis of statistical research either. Indeed, it should be the other way around—it seems more appropriate that future researchers and consumers of educational research should first be taught to understand the research process and the content of statistical research studies, and after that, delve into courses that require designing research and doing statistical analyses.

PROBLEMS IN COMPREHENDING EDUCATIONAL RESEARCH

The intelligent comprehension and use of published research in education requires two interrelated, but distinct, abilities: (1) the ability to comprehend exactly what is said in any research report and (2) the ability to evaluate the quality, merit or worth of written research reports—i.e., critical reading. Progress towards attainment of these abilities is the major goal of this text.

Confronted with the task of reading *Crime and Punishment* or any other Dostoyevsky novel, the average mature reader usually has sufficient vocabulary and prior knowledge of the nature of world events, human characteristics, and the schema of literature to be able to

comprehend the meaning and significance of the text. What does present a major hurdle for most English-speaking readers is the multitude of names Dostoyevsky uses to identify his characters. Sort out those names, and the mature reader will generally make good sense of Dostoyevsky.

A doctoral student new to reading research reports faces an even greater task than the reader of the Russian novel. The reader of a novel has a framework or schema for literature; s/he knows that there are characters that need to be understood, a problem that needs to be identified, a plan for resolving that problem, and the actual resolution of the problem. The beginning reader of research however, does not usually have a well-developed schema for research studies; therefore, just fitting the various bits and pieces of the report into an organized schema can be difficult.

Coupled to this unfamiliarity with the framework of research reports is limited prior knowledge of the subject of the report. An adult reader has a lifetime of knowledge about natural phenomena and human characteristics to bring to bear on a novel; but that same reader when reading research literature may not have the prior knowledge of the specific methods or findings of previous research the writer assumes most readers know.

But the major roadblock for readers new to the domain of **quantitative research** usually is the omnipresent statistic. Indeed, statistics are the defining trait of quantitative research. Writers use statistics to make the report more readable by reducing large amounts of data to more comprehensible terms, but to many readers, statistics only add to the muddle and confusion. The problems they present are numerous. First there are the terminology and symbols, to which the reader responds by saying "It's Greek to me" (which is not wholly irrational, as many of the terms and symbols are Greek). Besides novel terms and symbols is the bewildering array of numbers, all of which the reader knows are supposed to be significant (as the author has no doubt said that some of these numbers are significant to a .05 or .001 level). On top of numbers and terminology, the research report usually includes tables and graphs, which again may add to the difficulty of interpretation.

What makes it even more difficult for doctoral students is that they are literate and knowledgeable, and therefore sophisticated enough to understand that the writer has used these

statistical terms, symbols, statistics, and graphic presentations in order to make the presentation more concise and easier to comprehend; and since the reader fails to comprehend when knowing s/he should, what was a simple cognitive problem also becomes an affective problem. Unfortunately, the affective problems caused by the presence of statistics are far more difficult to overcome than the cognitive problems—i.e., the difficulties in comprehending reports containing statistics are more imagined than real.

A major goal of this text is to help readers develop **visual** and **verbal conceptions** of the statistics most often used in quantitative research. The focus on statistics is never in terms of calculation, derivation, or proof. Readers seeking formulas, derivations, and proofs should put this text aside and turn to any of several dozen excellent statistics texts. Instead, the focus here is on helping readers learn the premises, uses, and interpretations of certain statistics as they are commonly applied in educational research by helping them:

- understand the logic and necessity of statistics,
- **visualize** the groups and variables being statistically analyzed,
- **verbalize** the statistical analyses being made, and
- interpret the results of statistical analyses in **verbal** form (i.e., in written or spoken words) and in **visual** form (i.e., in tables or figures).

RATIONALE OF THE BOOK: STATISTICS AS VOCABULARY STUDY

Note that the tasks identified as important in learning to read statistical research include nothing that is mathematical. The study of statistics is not a mathematical challenge to graduate students—how can it be, the most complicated math in the first year or two of actual statistics is taking the square root of a number, all the rest is adding, subtracting, dividing, and multiplying. You mastered all these when you were in seventh grade—even if your grades in math were C's. No, math is not the obstacle to understanding statistics; the challenge in learning to read statistical research actually lies in the verbal domain, not the mathematical. The principal task in becoming a critical reader of statistical research is learning the **concepts** statisticians use and the **words** they ascribe to them.

Using the terminology of many philosophers, linguists, and psychologists who have pondered the learning of new concepts, I use the word *thing* instead of concept. The rationale used in this book for teaching things or concepts is that learning statistical things is verbal learning, and verbal learning is "fitting" or associating new things with old things—i.e., new knowledge with old knowledge or prior knowledge (Kibby, 1995). All learning—verbal, numerical, spatial—requires integrating new things with things already known. Learning is not just adding information on top of information known, but rather, it is integrating already-known things with new things. Each reader must construct his or her own conceptions. The two-time Nobel Laureate for Chemistry, Linus Pauling, vividly described the learning of new things:

> If I couldn't find a place for some thing, then I would change my picture of the world until I understood where it fit. Or, I would throw it out and come back to it later.

Learning to read statistical research is made complex not so much by the difficulty of the things to be learned (as will be explained in the next few paragraphs), but rather by the fact that there are so many things and words to learn. Indeed, when you complete this text, you will recognize that it is your vocabulary and conceptualization that has increased, not your math ability. The only math required in this entire book is the ability to average a group of numbers (in statistical jargon, called finding the mean of the scores); this statement is not hyperbole.

The degree of difficulty in learning of new things is largely dependent upon a learner's prior knowledge. Instruction in learning a new thing must, then, be varied according to the level of the learner's prior knowledge related to the thing to be learned. This book assumes that the things to be learned by doctoral students will fit one of four levels of prior knowledge.

Level 1—Production Knowledge: things students know and are able to associate with one or more words, but students must learn yet one more word for that thing. For example, learning that fingers and toes are also called "digits." Learning here involves focusing on the known concept, which can be easily done with a word(s), and then associating that known thing with a new word. The word "mean" (which means the same thing as "average") is Level 1 learning for most students needing to read this book.

Level 2—Organized Prior Knowledge: things students know but cannot associate with any word or term for the thing. For example, we all know that there are two little lines between the nose and upper-lip—that is, we know the thing—but very few know that this thing has a word, "philtrum." Thus, learning this word is a matter of focusing attention on the thing that is already known—by description or pictures—and then associating that known thing to the new word. The symbol μ (Greek letter m, called mu) and the term "alternative hypothesis" are likely examples of Level 2 learning for students using this text.

Level 3—Immediately Learnable Knowledge: students do not know the thing they must learn, nor do they know the word for that thing, but they have sufficient prior knowledge to conceptualize or visualize the thing with verbal or graphic descriptions or definitions. For example, most adults can conceptualize an original oil painting being painted over by a second oil painting, and also visualize parts of that second layer ebbing away until the original painting can be seen, but most adults do not know that what can be seen in that lower layer is called "pentimento." In this case, learning the term pentimento, as Pauling described, means learning to reorganize or reconceptualize knowledge one already has (i.e., old painting, new painting, wearing away of old painting) by description or pictures and associating that new conception or thing with a word that also is new. A good proportion of statistical things are Level 3 learning, including "probability distributions" and "SE_m" (standard error of the mean).

Level 4—Potentially Learnable Knowledge: students do not know the concept or thing that they must learn, nor do they have a word for it. Take the word and concept "osmosis" for example. When we learned this as grade- or middle-school students, we could already conceive a greater concentration of liquid and a lesser concentration of liquid. Thus learning about the movement of a liquid from a greater to lesser concentration was not the difficulty in learning osmosis: the difficulty was learning the thing called a "semi-permeable membrane," and that was the chief teaching task. In learning statistics, the words "kurtosis" and "skew" are most likely to be Level 4 learning.

What makes this conceptualization of four levels of prior knowledge for learning a new thing important here is that a great many of the things that statisticians use are already a part of your production, organization, or immediately learnable knowledge—that is already known or partially-known. What you need to learn is to hook or associate this known or partially-known thing to the unknown term, word, or symbol statisticians use to identify the thing. For example, the idea of comparing three groups after experimental treatment x has been given to one, experimental treatment y has been given to the second, and nothing has been done to the third is an idea that every adult who has watched or read toothpaste advertisements already understands. But when a statistical researcher describes this simple process with terms such as stratified random sample, random assignment, means, standard deviations, univariate 1-Way ANOVA, *F*-ratio, significant, post hoc analyses, orthogonal contrasts, degrees of freedom, attrition or mortality, effect size, etc., the poor reader is lost—and most often says, "Math has always been hard for me." The problem here, however, is definitely not math, the problem is verbal—the meaning and usage of terms, symbols, and words. Moreover, like most graduate students, you have long mastered the intricacies of learning new words, thus many of the terms in the preceding list are things you will be able to learn quite readily—though in statistical research the number of new terms can be nearly overwhelming. This book works on the premise that learning statistics is largely learning the jargon of statistics. Once you become convinced of this, most (never all) of the anxiety about statistics fades and learning becomes easier.

Let me give a concrete example of this major point. Perhaps the most important thing for you to learn in order to read statistical research is the "standard error of the mean," or, SE_M. Now this is a darn difficult concept for students to learn, and I am not trying to say that learning it is easy. But it is made less difficult to learn if instruction shows you how you already know most of the things related to this term. If you are among the typical American adults who drink Diet Coke© with your fudge brownies, you are aware that if you weigh yourself every day, you will not weigh exactly the same each time. You also know that without dieting and exercising or overeating, that this variation will not be very large (depending upon your weight). Indeed, you could probably

sketch a fictional line graph of what that variation might look like from your personal knowledge of your weight fluctuations. Such life-experiences are used throughout this book to demonstrate statistical logic or terminology, and I believe it goes a long way to demystifying statistics.

Relying on several years experience teaching students to read statistical research, I try in this book to identify the likely level of difficulty (1, 2, 3, or 4) of the thing to be learned, use readily understood examples from daily living, help you discriminate or isolate the important things of what it is you already know, and focus your attention on these points by generous use of graphics.

To be sure, there are statistical things you must learn that are not already a potentially-distinctive thing in your prior knowledge, things that are only potentially learnable knowledge at the start of instruction (Level 4 difficulty). In teaching these, I try to provide well thought-out written descriptions and carefully-designed graphic representations. Included among these potentially learnable concepts for most doctoral students are "degrees of freedom," "interaction effects," "main effects," "regression to the mean," "kurtosis," "skew," "curvilinear relation," and "orthogonal contrasts." Your learning these is done best, however, on the rationale that verbal learning of vocabulary, terminology, or symbols is the major learning task, not the mastery of mathematical knowledge or procedures.

WHAT IS RESEARCH AND EDUCATIONAL RESEARCH

In 1897, T. C. Chamberlin—a geologist and one of the fathers (with F. Moulton) of the planetesimal theory of the creation of the earth—stated that "there is no nobler aspiration of the human intellect than the desire to compass the causes of things" (p. 842). He defined research as "primary or creative study . . . to discover new truth, or to make a new combination of truth, or at the least to develop by one's own effort an individualized assemblage of truth" (p. 837). Before and after Chamberlin, numerous essayists opined on the nature and definition of research and the various terms by which research is known—e.g., disciplined inquiry (Cronbach & Suppes, 1969),

inquiry (Eisner, 1991; Shulman, 1981), scientific research (Kerlinger, 1986; Platt, 1966), basic and applied research (Ausubel, 1953), naturalistic research (Lincoln & Guba, 1985).

Research is a process of systematic or disciplined inquiry conducted for the purpose of addressing a narrowly-defined question, almost always a question about the association of two or more constructs, traits, or variables. In a frequently reprinted section of the Cronbach and Suppes book (1977), the term *disciplined inquiry* is used to identify research that has been well conceived, conducted, and reported. These critical processes of disciplined inquiry apply to any genre of research—experimental, correlational, descriptive, ethnographic, case study, historical, reviews, critical, evaluation, or assessment—and such disciplined inquiry:

- anticipates traditional questions,
- institutes controls at each step of information or data gathering,
- employs reasoning to avoid sources of error,
- takes errors into account by discussing margin for error in conclusions, and
- displays the raw materials of an argument as well as the "logical processes by which they were compressed and rearranged to make a logical conclusion" (pp. 15-16); Cronbach and Suppes expand this presentation of the argument in disciplined inquiry:

  > The detail of the argument, whether it is describing methods of data collection or the derivation of practical recommendations, is lucid, specific, and pertinent. With such a presentation there is something to be learned from *explication de texte*, whereas in an undisciplined discussion the summary message is all that can be taken seriously.

What differentiates research in education from research in other fields are not research designs, methods of inquiry, or discipline, but the questions addressed by the inquiry and the uses made of the results. It would be too restrictive to define educational research as investigations of the conditions for teaching certain information via schoolrooms, teachers, and materials, as the full spectrum of research in education is not restricted to direct teaching, nor to classrooms. Educational research is much broader; it encompasses all investigations of the conditions of instruction intended to transmit specified knowledge and understanding for the purposes of

deliberately influencing learners for some greater social good. The three purposes of educational research are (a) to enhance the theory or knowledge of the conditions of schooling, teaching, and learning; (b) to lead directly to the improvement of teaching or learning; or (c) to lead to the improvement in methods or materials of educational research.

Educational research is concerned with the gamut of issues related to learning—what is to be learned and why, when and where it will be learned, by whom it will be taught, and how learning is to be evaluated. Certainly other academic disciplines share a keen interest in the subject of *learning*. But other disciplines usually invest their research energy into only one or two of the many components related to learning as conceived by educators. Psychologists study learning, but their aim is generally to understand the psychological or physiological phenomena that accompany, limit or facilitate learning or processing a very precisely-defined form of content in tightly-controlled environments. Educators have profited immensely from the work of learning psychologists, but they first had to conduct (or, in too many instances, should have conducted) a great deal more research using expanded learning tasks with normal students in ecologically valid environments (Bronfenbrenner, 1977) before the learning psychologists' theories could be applied to education. Sociologists may also study learning and how learning is influenced by sociological factors such as socioeconomic status or group membership. Education has profited greatly from sociological research, indeed, it literally transformed funding for educating the poor and disadvantaged in the 1960s. But questions of what knowledge should be learned, how that knowledge should be organized and transmitted, and how to evaluate the effects of the transmission of this knowledge are questions that are not addressed, rightfully so, by sociologists. Philosophers and linguists may devote large portions of their lives to studying the acquisition and organization of language and knowledge, and certainly the fruits of their labor inform educators in a significant manner. But philosophers and linguists do not generally consider the important questions of the formal transmission of language and ideas in school settings.

THE TERMS *RESEARCH* AND *INQUIRY*

Cronbach and Suppes (1969) observe that "too many writers seem to limit the term 'research' to quantitative empirical inquiry" (p. 14), which they think gives short shrift to other genres of research, especially qualitative and historical inquiry; this leads them to prefer the term *inquiry* instead of *research.* They believe that "the study of education requires non-quantitative as well as quantitative techniques" (p. 14).[1] This text argues that all forms of research are necessary in education, and no single research genre will ever provide a full solution to any significant educational question—however, the focus of this book is statistical research. The terms *research* and *inquiry* are synonymous here, and they apply equally to any form of educational research.

A METACOGNITIVE ORGANIZER OF EDUCATIONAL RESEARCH

BASIC, APPLIED, AND ACTION RESEARCH

The major components of basic and applied research are the same: they address significant questions by systematically collecting information, analyzing or reducing that information in a credible manner, and drawing conclusions from the methods and findings. If the purpose of the research is to discover general laws of certain phenomena and there is no immediate social benefit to the findings and conclusions, we generally call the research basic; if the findings and conclusions have a **valid** and immediate application (or point to a possibly **valid** solution) to some socially significant issue, then we usually call the research applied. The key word here is valid. It is neither interesting or important to debate if a given educational study is basic or applied research, what is important is to determine if the findings of the study can be validly applied to a current educational dilemma. That decision is determined less by the findings of the study then by the

---

[1]Erickson (1985) would not likely argue with Cronbach and Suppes' point that the study of education requires both quantitative and qualitative research, but he would argue that qualitative research is and should not necessarily be "non-quantitative"; he states that "quantification of particular sorts can often be employed in" qualitative methods (p. 119). Indeed, he prefers the term interpretive research over qualitative research.

methods and subjects of the study. A study that finds pigeons respond to a buzzer by pushing a lever when they are rewarded with pigeon food may provide specific knowledge on the conditions of one kind of learning and reinforcement of that kind of learning—but only for pigeons. To use the findings of such a study as justification for giving M&Ms as rewards for correctly solving multiplication problems is a huge and invalid leap. When basic research in learning is conducted, it is usually conducted in a laboratory or tightly-controlled environment with very highly-selected cases or subjects who are all required to learn precisely-defined types of content under specified instruction or exposure. None of these conditions is found in even the most ideal (or bizarre) learning environments, thus to apply the findings of such basic research to schools and classrooms without further research—applied research—is, in spite of its frequent occurrence, foolhardy. Ausubel (1953) uses the term "naiveté" to describe the direct application of basic research to real educational milieus; he is perhaps too polite.

Action research in education is usually conducted by persons directly confronted with an educational problem for which they find no solution in published literature or via consultation. In education, regular classroom teachers or first-line administrators (team leaders, coordinators or principals) are those most likely to conduct action research, though they often seek consultation from researchers on design, data collection, and data analysis issues. The usual purpose of action research is to solve a concrete problem with a known group of students already present in the teacher's environment. The research is not designed with the intention of addressing the problem in a manner in which the findings can be validly applied or generalized to persons other than the actual participants in the study, though it is possible the results may have such generalizability.

CLASSIFICATION OF THE DOMAINS AND GENRES OF EDUCATIONAL RESEARCH

All research addresses a research question, problem or hypothesis by the organized collection, examination, and interpretation of information—i.e., data. But researchers vary in their general philosophies or conceptualizations of the world and its inhabitants, in the nature of the research problems they address, in the information or data they seek, and in the methods they employ for

analyzing and interpreting data. These differences result in a division of the full range of educational research into categories—in this book, four major categories which are called *domains*. In addition, within each domain there are sub-groups of research types, that differ further in terms of research goals and the analysis and interpretation of data. In this text, these sub-groups are called *genres*, and there are 10 of them.[2] Knowledge of these four domains and ten genres also provides a metacognitive structure to the field of educational research. This text does not cover all four domains and ten genres—any text attempting to do a credible job of this would be so long as to be burdensome (trust me on this, for when this book was started, it was intended to cover all 10 genres). But those beginning the journey of learning to read educational research should develop a metacognitive overview or gestalt of that field.

The first major step in reading educational research is to recognize its specific genre, and then to apply knowledge of that genre of research to the specific research report. This makes you a more active reader allowing you to use prior knowledge to anticipate the general structure of the report and to look for specific information within it. The four domains and ten genres are:

Domain A: Quantitative Research

Genre I: Experimental Research

Genre II: Correlation and Regression Research

Genre III: Quantitative Descriptive Research

Domain B: Qualitative or Naturalistic Research

Genre IV: Qualitative Descriptive Research

Genre V: Case Study Research

Domain C: Document Analysis Research

Genre VI: Historical Research

Genre VII: Reviews of Research

Genre VIII: Analytical, Theoretical and Critical Research

Domain D: Program and Policy Research

Genre IX: Program Evaluation Research

Genre X: Assessment Research.

---

[2]The word "paradigm" is commonly used to indicate various strategies for conceiving and designing research; but because that word has been involved in such heated debates in the last decade, I have opted to use the word genre.

There is very little that you will be asked to memorize in this book, but you should now commit to memory these domains and genres, even though this book will not cover Genres IV, V, VI, or VIII.

FURTHER CLASSIFICATION BY RESEARCH DESIGNS AND DATA ANALYSIS TECHNIQUES

For most readers, the three most difficult components of research reports are the (a) research design, (b) methodology, and (c) data-analysis techniques. Indeed, a few doctoral students have been overheard to say that they "just skip these parts and read the introduction and conclusions." Such a strategy hardly leads to critical reading of research. Although coverage is given to nearly all terminology and techniques pertinent to statistical research, an especial focus will be given here to designs, methods, and data-analysis techniques. After completing this text, readers should not have to "skip" these technical aspects of research reports, as they should be able:

1. to comprehend, sketch and visualize the research **designs** reported in published research and determine if those designs are appropriate for the research problem;
2. to apprehend and visualize the research **methodology** and evaluate the appropriateness of this methodology for the given research problem; and
3. to comprehend the logic of the **data-analysis** techniques; predict, follow, and visualize the specific contrasts or analyses made; comprehend and interpret (verbally and graphically) the results of these data analyses, and evaluate the appropriateness of the author's data analyses and interpretations.

Indeed, such attention is paid to these technical issues, that after first classifying the breadth of educational research by domain and genre, this text applies a second level of classification: within each genre, studies are classified by their (a) research design and (b) major data-analysis technique. Within the Experimental and Correlation/Regression genres, this second level of classification also orders research by level of difficulty; thus easier research designs and data-analysis techniques are encountered before difficult ones. Figure 1-1 displays the four major domains of research in the left column and the genres within each domain serve as the main entries in the right column

(domains and genres are presented in boldface type). Indented under the names of the genres are the major categories of research designs and data-analysis techniques of that genre (normal typeface).

Insert Figure 1-1 About Here

BRIEF DESCRIPTIONS OF THE DOMAINS AND GENRES

The following section provides brief descriptions of the goals, methods, and data-analysis techniques of the ten educational research genres. Obviously, the remainder of this book will provide in-depth coverage of the statistical-research genres (Genres I, II, III, VII, IX, and X), but the following section provides a brief overview of all domains and genres, and is intended to help further your ability to construct a meta-cognitive organization of the realm of educational research.

Quantitative research (a) uses quantitative measurement instruments or tests to collect numerical data about predetermined, specified traits or characteristics from a number of cases (persons, materials, methods), (b) statistically reduces these large amounts of data into more succinct numerical forms (e.g., averages, measures of dispersion), and (c) then uses additional statistics to analyze the data for trends, associations, or differences among the traits or cases. In experimental research (**Genre I**), the researcher directly varies some trait or characteristic (called the independent variable) of persons, materials, methods, etc. in order to determine the effects of that variation on some other trait(s) or characteristic(s) (called an outcome measure or dependent variable). Those effects are almost always measured in some quantitative manner. An example of experimental research is when a researcher compares the rate of tooth decay of one group of persons who brushed their teeth for a year with a toothpaste containing compound x to another group who brushed their teeth for a year with toothpaste without compound x. Most experimental studies are based on historically well-documented research designs (Lindquist, 1954; Campbell & Stanley, 1963), generally *t*-test, several forms of analysis of variance (ANOVA), and nonparametric techniques. These designs, which are also data-analysis techniques, are shown in Figure 1-1 and will be covered in this book.

In correlation and regression research (**Genre II**), the researcher usually does not manipulate or try to change the subjects in any manner (though such is possible), but instead, generally measures quantitatively two or more traits of a sample of cases, and then quantifies the relation between these measures. For example, if a researcher quantifies how often each person in a large group of persons brushes his or her teeth (one trait or variable) and how many fillings each person has (the second trait or variable), the researcher can use a correlation coefficient to determine the quantitative degree of association between these two variables. Regression is an extension of correlation. In using regression statistics, a researcher may predict a person's or group of persons' score on one trait from their scores on one, two, three or more other traits. As a college student you knew the concept (thing) regression, you just did not know that this is what we call it. On the basis of previous regression research, colleges and universities are able to predict (with a known margin of error) how well students will do in their first year of college on the basis of those student's SAT or ACT scores and high school grades. Analysis of covariance (ANCOVA) is yet one more extension of correlation. A researcher may compare two (or more) experimental treatments on groups that were not equal at the beginning of the study, and adjust the scores of each group on the measure administered at the end of the study for the effects of the differences at the beginning of the study. Factor analysis is used to examine large data sets in which the same set of cases is measured on a large number of traits. Researchers first set up large correlation matrices in which level of performance on every trait is correlated with level of performance on every other trait. Factor analysis then examines these huge matrices in order to find groups of traits that are much more highly associated with each other than with other traits. Intelligence as conceived in the last 60 years is a good example of factor analysis. By examining huge matrices of correlations between scores on a large variety of tests designed to measure intelligence, Thurstone (1938) identified several fundamental factors that underlie intelligence, and the names he gave these factors are today used in ordinary conversation (e.g., verbal ability, numerical ability, reasoning, spatial ability). To be sure, scores on tests of verbal ability do correlate with scores on tests of numerical ability, but each of the tests of verbal ability correlates more highly with other tests of verbal ability

than with tests of numerical ability, and the reverse is the case for tests of numerical ability. Correlation, six varieties of regression, and brief introductions to ANCOVA and factor analysis are listed in Figure 1-1, and each will be developed in this book.

This text classifies quantitative-descriptive research **(Genre III)** into five different types of studies, but they all address the general purpose of providing a quantitative description of status quo. Data are collected on one or more important variables or traits of persons just as those persons are (that is, they are not exposed to experimental adjustments or treatments), and these data are quantified in some manner. One form of quantitative-descriptive research is the causal comparative study (sometimes called ex post facto). Such studies compare two (or more) groups known to differ on one important variable(s) on one or more other variables. For example, a study comparing the academic performance of children whose parents are alcoholics to children whose parents are not is a measure of the status quo of two groups. This study looks as if it might be an experiment, but the two groups of children were exposed/not exposed to alcoholism in their natural environment, and no experimental effect was put into place to cause certain parents to become alcoholics. Indeed, such an experimental effect, though fathomable, is unconscionable. A second form of the quantitative-descriptive research study is the developmental study, and its purpose is to determine changes over time. It has two forms: longitudinal, in which one or more groups of persons are observed over a long period, and cross-sectional, in which several groups of persons varying in age (or some other developmental scale) are compared at the same time. Surveys are the third variety of quantitative-descriptive research, such as a Gallup Poll that reports how many people will vote for candidates x and y. A fourth example is the content analysis. For example, Carroll, Davis and Richman's (1971) analysis of 1,000,000-plus words from all varieties of print (e.g., magazines, books, newspapers, newsletters, etc.) for the purpose of listing every different word and how many times each word appeared. The last category of quantitative-descriptive research is the a priori observation study. In these studies, the researcher determines before going into the classroom, etc. exactly what behaviors, etc. s/he will observe (compare this process to the qualitative researcher's approach described in the next section). An example of a priori research is

Durkin's (1978-79) oft-quoted study of reading comprehension instruction in grades four, five, and six. In her paper, she spends several pages describing what she meant by reading comprehension and attached a multi-page appendix of activities that could be counted as instruction in reading comprehension or other forms of instruction or non-instruction. Durkin's trained her observers to use this list of instructional and learning behaviors, who then went into classrooms and tallied the number of times they observed the given activity.

Qualitative descriptive research, often called **naturalistic research**, (a) collects data about certain traits or characteristics—which are not always defined before the research starts—about the cases or participants by observation in natural settings, and records those data by either written notes or audio/videotape, (b) analyzes and reduces these notes or recordings in a non-numerical manner (e.g., constant-comparative) and (c) and reports the data almost entirely in words, not numbers or statistics. The goal of qualitative descriptive research (**Genre IV**), like quantitative descriptive research, is to provide a description of status quo, but that description is for the most part verbal (which means either written or oral), and statistical or any other quantitative data usually play only a minimal role. Qualitative research is theory driven, but the qualitative researcher often does not formulate his or her theories or hypotheses until the researcher has been in the *field* a while—the field is the natural environment of the cases or situations that the researcher is studying. The researcher enters the field of the study with as much information as possible, but tries gallantly to delay making assumptions or hypotheses before putting in hours of observation or interviews. In the field, the researcher may either stand off to the side and observe (non-participant observation) or jump into the thick of things and interact completely with the participants in the milieu (participant observations). In either form of observation, the researcher spends a great deal of time (weeks, months) in a specific environment trying to understand that environment and its participants—from the point of view of the participants. The researcher grounds his or her theories and hypotheses in what s/he observes, hears, and reads, not on the basis of a priori conceptions (Glaser & Strauss, 1967). Such research is based upon the model established by early 20th century anthropologists. An excellent example of participant observation

research in education is Alan Peshkin's (1986) *Gods Choice*. Peshkin spend almost two full years in a Christian school, going to classes, PTA meetings, sports activities, staff meetings, homes of teachers and students, church, prayer meetings, and many other activities associated with the school. The purpose of his book was to describe, from the point of view of the parents, teachers, and students in the school, what life was like attending a Christian school. Interview research is based on interviews—usually very long and occurring over an extended period of time.

Case study research (**Genre V**) is also descriptive, and some of it may use quantitative data extensively (Robinson, 1946), but generally uses those quantitative data in conjunction with many other sources of non-quantitative data. The purpose of case study research is to describe one person or one group of persons sharing a single underlying and important trait (e.g., alcoholics, disabled readers, twins, gifted students) in as much depth as possible. The descriptions resulting from both Genres IV and V research usually lead to a better understanding of certain groups, milieus, or organizations; new theories that need further study and exploration; or generalizations that might apply to other similar groups, milieus, or organizations (though the latter point is much debated). One of the most influential studies in the history of reading was a case study conducted by Helen M. Robinson in the late 30s and published in 1946. Robinson studied 30 children with reading problems. Each student was evaluated by a team of physicians including a neurologist, ophthalmologist, otolaryngologist, pediatrician, and endocrinologist. Each child was also evaluated by a speech and language specialist, a clinical psychologist, a social worker, and educational specialists. In no case was the cause of a child's reading difficulty unequivocally established. Robinson concluded, that with the exception of children with severe sensory deficits (e.g., deaf, blind, developmentally disabled), it is impossible to determine positively the cause of a child's difficulty in reading. Since Robinson's study, no legitimate reading diagnosis has operated on the premise that the purpose of reading diagnosis is to find the cause of the reading problem. Though an important and growing field of research in education, this text will not cover the qualitative domain.

Document analysis research describes research that uses extant texts/materials as their main source of data. Document analysis research shares traits of quantitative and qualitative research in that it (a) collects either quantitative or verbal data from extant texts and materials (b) reduces those data either quantitatively or qualitatively, and (c) analyzes the data either statistically or in a non-quantitative analytical manner. Historical research (**Genre VI**) uses previous information (published and unpublished reports, archival information, logs, journals, receipts, minutes of meetings, etc.) for the purpose of describing, analyzing, or interpreting past events or persons. One purpose of historical research is simply to know what happened in the past, another purpose is to learn how certain problems confronting peoples in the past were addressed (Gottschalk, 1954). Although this is the form of research I most like to read for pleasure, this Genre will not be covered in this book.

Research reviews (**Genre VII**) summarize and interpret past research, usually experimental, quasi-experimental, ex post facto, or correlation/regression research, for the purposes of summing up what is known about a given problem and drawing generalizations from that body of research. For example, Holmes (1989) reviewed all published and unpublished research on the effects on achievement and affect of retaining—flunking—students in the primary grades. Traditional or ballot reviews (sometimes called bullet reviews) review all the studies that have examined a given experimental or quasi-experimental effect, and tally the number that found and did not find significant results. Meta analyses apply a statistical procedure to data from these studies to obtain a quantitative estimate of the effects of the treatments. A third form of research reviews, perhaps the most interesting to read, is the essay or "think piece" in which the writer demonstrates a wide-ranging knowledge of the research literature (by citations and references), but it is clear that the writer is using this literature base to make a point, not necessarily to review the entire field from a totally objective point of view. All three forms of reviews will be addressed in this book.

There are three major forms of criticism (**Genre VIII**): the critical review, critical research, and theory or model building. Research criticism usually consists of (a) an original

research study (of any genre, including research reviews), (b) a critical essay or analysis of that study, and (c) a rebuttal of the critique by the author of the original study. Usually the study that is critiqued is significant, and often the review is written by someone who offers a different theoretical interpretation of the data proffered in the original study. The 1959 review by Noam Chomsky (a cognitivist) of *Verbal Learning* (1957) by B. F. Skinner (a behaviorist) was important because of Chomsky's differing theoretical interpretations of Skinner's data. This category does not include simple critiques that merely dismiss a piece of research for lack of quality. The second from of criticism is critical research, which might also be called philosophical research. Here, research and analysis are brought together in a data-based essay, an essay which usually addresses a socially or educationally significant issue. Inherent in the argument of this form of research is that some widely-accepted value, method, material, or curriculum needs to be eliminated or altered. All research is either theory based or leads to the development of theory or models, thus a special category set aside for theoretical research might seem redundant. Further, all theoretical research actually utilizes the historical, review, and criticism genres of research, again indicating possible redundancy. The previously-mentioned Chomsky review certainly pointed to alternative theoretical conceptualizations of verbal learning, and if it had not been written as an invited review of Skinner's book, could well be placed in this category. What makes theoretical research a form of its own is that some research—though built on the backs of other research—is conceived, written, and published only for the purpose of bringing new or alternative theoretical conceptualizations or models to research problems that have long confronted us.

Program and policy research (a) for the most part, collects quantitative data from cases about specified traits (qualitative data may also be collected, though this is less common), (b) reduces these data statistically, and (c) using additional statistics, analyzes these data. Experimental, correlational, regression, quantitative, case study, historical, or research reviews may be used as part of their basic methodology of program and policy research. However, evaluation research (**Genre IX**) is decision oriented, and is conducted to inform the researcher on the costs, effects, and worth (value) of specified goals, methods, materials, curricula, staff,

facilities, etc. A study to determine the benefits of a one-to-one remedial reading program for at-risk first graders would not only examine the effects of this program on achievement and future achievement, but would also report the costs of this program in terms of dollars, staff time, facilities, materials, etc. The main goal of assessment research (**Genre X**) is to report the current performance or achievement status of certain cases (e.g., fourth graders) on a specified trait(s). For example, through the auspices of the National Center for Educational Statistics, every few years the National Assessment of Educational Progress (NAEP) administers reading, math, writing, and other academic tests to representative samples of American fourth, seventh, and eleventh graders in order to determine the current status of reading, math, etc. and to ascertain changes (trends) in achievement over time.

General Organizational Structure and Quality Standards of Educational Research Reports

To become a critical reader of any given educational research report, a reader must have developed at least three knowledge bases (beyond prior knowledge of the topic of that report):

1. knowledge of the report's organizational structure—to become an active reader and be able to anticipate text, to know when important information has been missed or not included in the report, to be able to organize and retain information as it is presented;
2. knowledge of the terminology and symbols used in the report—to understand exactly what the researcher is saying, calculating, or finding; and
3. knowledge of the standards by which to judge the major elements of research reports of this genre—to be able to determine the report's quality, merit and contribution.

Terminology, organization, and quality standards vary from genre to genre, and in the course of this text, those terms, organizations and quality standards germane to each genre will be explained and demonstrated. However, the differences between genres in the style and organization of reports are all variations on a general structure of all research reports. As a first step

toward becoming a critical reader of research, it is important that readers learn the general organizational structure of educational research reports as displayed in Figure 1-2.

Insert Figure 1-2 About Here

In a manner similar to style and organization, the differences between genres in standards of quality are also variations on a general set of expectations of high-quality educational research. Taking into account Cronbach and Suppes' (1969) guidelines and adding a bit more specificity, Figure 1-3 presents a set of the elements of a research study to be considered in evaluating its quality. As the text moves into each specific genre of research, more precisely honed elements and questions will be presented; more importantly however, the text will present the guidelines and standards to be used in responding to these elements or questions.

Insert Figure 1-3 About Here

BECOMING LITERATE IN EDUCATIONAL RESEARCH

Scholarship means cognizance of the current and historical theory, knowledge, methods, materials, equipment, and applications of a specific academic discipline, and maybe even occasionally contributing to that field of knowledge; thus scholarship requires continual reading—critical reading—of the spectrum of published and unpublished research of the discipline. To keep up with the ever-expanding knowledge base of intersecting areas within the discipline of education, even the most die-hard experimentalist must be able to read qualitative and historical research critically, and at the same time, the died-in-the-wool ethnographer has to be able to understand what is being reported in quantitative research articles. There is no longer room in education and the social sciences for bickering between quantitative positivists and qualitative constructivists as to whose view is correct; those with any sense have come to recognize that both ends of the quantitative-qualitative continuum are necessary—and that no single research genre will ever provide a full solution to any significant educational issue. The fact that this book is devoted only to quantitative research in no way gainsays other forms of research, but is dictated by exigencies such as realistic book lengths and the amount of content that can be covered in one course.

Some educators in leadership positions in elementary and secondary schools may believe that educational research is not important to them in their day-to-day activities of administering, coordinating, or evaluating programs and persons. Nothing could be further from the truth. An educational practitioner will be called upon time and again to select, design, and evaluate new materials, new curricula, and new methods; and in these situations, the practitioner should have the abilities necessary to read critically the research reports by others who have investigated these new materials, curricula, or methods. A reader who cannot critically evaluate research, including the technical aspects of a study (e.g., design, subjects, procedures, materials, measurement instruments, and data-analysis techniques), will have to trust that that study's findings and conclusions are valid, reliable, and generalizable to that practitioner's needs and situation. Decisions as important (and, often, as expensive) as these should not be left to "trust."

REFERENCES

Ausubel, David P. (1953). The nature of educational research. *Educational Theory, 3*(4), 314–320.

Baldwin, R. Scott, Readance, John E., Schumm, Jeanne Shay, Konopak, John P., Konopak, Bonnie C., and Klinger, Janette K. (1992). Forty years of NRC publications: 1952-1991. *Journal of Reading Behavior, 24*(4), 505-532.

Bronfenbrenner, Urie. (1977). Toward an experimental ecology of human development. *American Psychologist, 32*, 513-531.

Campbell, Donald T. & Stanley, Julian C. (1966). *Experimental and quasi-experimental designs for research.* Chicago: Rand McNally.

Carroll, John B., Davies, Peter, & Richman, Barry. (1971). *The American Heritage word frequency book.* New York: American Heritage.

Chamberlin, T. C. (1897). Studies for students: The method of multiple working hypotheses. *Journal of Geology, 5*, 837–848.

Cronbach, Lee J., & Suppes, Patrick J. (1969). *Research for tomorrow's schools.* The report of the committee on educational research of the National Education Academy. New York: Macmillan.

Durkin, Dolores. (1978–1979). What classroom observations reveal about reading comprehension instruction. *Reading Research Quarterly, 14*(4), 481–534.

Eisner, Elliot, W. (1991). *Qualitative inquiry and the enhancement of educational practice.* New York: Macmillan.

Eisner, Elliot W., & Peshkin, Alan. (1990). *Qualitative inquiry in education: The continuing debate*. New York: Teachers College Press.

Elmore, Patricia B., & Woehlke, Paula L. (1988). Statistical methods employed in *American Educational Research Journal, Educational Researcher*, and *Review of Educational Research* from 1978 to 1987, *Educational Researcher, 17*(9), 19-20.

Erickson, Frederick. (1985). Qualitative methods in research on teaching. In Merlin C. Wittrock (Ed.), *Handbook of research on teaching* (3rd Ed.) (Chapter 5). New York: Macmillan.

Glaser, Barney R. & Straus, Anselm L. (1967). *The discovery of grounded theory: Strategies for qualitative research.* Hawthorne, NY: Aldine de Gruyter.

Glesne, Corrine, & Peshkin, Alan. (1992). *Becoming qualitative researchers: An introduction.* White Plains, NY: Longman.

Gottschalk, Louis R. (1954). *Understanding history.* New York: Knopf.

Holmes, C. Thomas. (1989). Grade level retention effects: A meta-analysis of research studies. In Lorrie A. Shepard & Mary Lee Smith (Eds.). *Flunking grades: Research and policies on retention.* London: Falmer Press.

Kerlinger, F. N. (1986). *Foundations of behavioral research* (3rd Ed.). Fort Worth, TX: Holt, Rinehart and Winston.

Kibby, Michael W. (1995). The organization and teaching of things and the words that signify them. *Journal of Adolescent and Adult Literacy, 39*(3), 208-223.

Light, Richard J., & Pillemer, David B. (1984). *Summing up: The science of reviewing research.* Cambridge, MA: Harvard University Press.

Lincoln, Yvonna S., & Guba, Egon G. (1985). *Naturalistic inquiry.* Beverly Hills, CA: Sage Publications Inc.

Lindquist, E. F. (1954). *Design and analysis of experiment in psychology and education.* Boston: Houghton Mifflin.

Peshkin, Alan. (1986). *God's choice.* Chicago: The University of Chicago Press.

Platt, John R. (1964). Strong inference. *Science, 146,* 347–352.

Pugh, A. K. (1984). The relevance of the study of the history of reading. In Greg Brooks and A. K. Pugh (Eds.), *Studies in the history of reading.* Reading, England: Centre for the Teaching of Reading, University of Reading School of Education with the United Kingdom Reading Association.

Robinson, Helen M. (1946). *Why pupils fail in reading*. Chicago: The University of Chicago Press.

Shulman, Lee S. (1981). Disciplines of inquiry in education: An overview. *Educational Researcher, 10*(6), 5–12, 23.

Smith, Mary Lee, & Glass, Gene V. (1987). *Research and evaluation in education and the social sciences*. Englewood Cliffs, NJ: Prentice-Hall.

Thurstone, L. L. (1938). *Primary mental abilities*. Chicago: The University of Chicago Press.

Figure 1-1. System for classifying the varieties of educational research by domains/genres and designs/data analysis techniques.

Domain A: Quantitative Research	**Genre I. Experimental** *t*-tests 1-way ANOVA ANOVA with repeated measures 2-way ANOVA 3-, 4- & 5-way ANOVA MANOVA Nonparametric methods (χ^2, *U*, Wilcoxon *T*, Sign Test) **Genre II. Correlation and Regression** Correlation Regression and multiple regression Linear regression Multiple regression Step-wise regression Regression using sets of variables Hierarchical regression Model testing using regression ANCOVA and Factor Analysis **Genre III. Quantitative Descriptive Research** Causal-comparative (quasi-experimental / ex post facto research) Developmental (longitudinal and cross-sectional) Survey research Content analyses A priori observations
Domain B: Qualitative or Naturalistic Research	**Genre IV. Qualitative Descriptive Research** Non-participant Observation Participant Observation Ethnography Community Studies Ethnography of Communication Field Studies Phenomenological Studies Interview Research **Genre V. Case Study Research**
Domain C: Document Analysis Research	**Genre VI. Historical Research** Historical research **Genre VII. Reviews of Research** Traditional or ballot reviews Meta-analysis Interpretive reviews and essays ("thinkpieces") based on research reviews **Genre VIII. Analytical, Theoretical and Critical Research** Critical reviews Critical research Theory / Model building
Domain D: Program and Policy Research	**Genre IX. Program Evaluation Research** Quantitative evaluation Qualitative evaluation **Genre X. Assessment Research**

Figure 1-2. The four major general organizational units of educational research reports and a statement and description of the major components within each unit.

	1. Purpose of the Research Study (The Research Problem)
Purpose	A statement of the purpose of the research, which may take the form of a question or hypothesis.
Background	No research problem or hypothesis is context free. The background of the research problem places that problem within the context of prior research, theory, or logical analyses of observations.
Significance	Statements as to why this research problem is significant: i.e., why it is necessary to have the answer to the research question or problem.
Need	Assurance that the researcher is familiar with the extant literature on this topic and knows that this research problem cannot be answered on the basis of other studies.
	2. Design and Methodology
Design	A description of the general design to be followed to collect observations or data.
Cases	A description of the persons, objects or actions to whom the experimental treatments will be applied or that will be observed or measured.
Specification of Procedures	An almost step-by-step specification of how the persons, objects or actions will be observed or measured or how the experimental treatments will be applied.
Outcome Measures	A precise statement of what data will be recorded from the observations, measurements or experimental treatments (sometimes included as a first component of the results section).
	3. Results
Summary Data	A listing, tabulation, description, or some other form of presenting a reduced form of the data obtained from the observations, measurements or experimental treatments.
Data Analysis Techniques	The general quantitative or qualitative data-analysis techniques that will be used to analyze the data derived from the observations, measurements or experimental treatments.
Findings	The specific results of the data analyses.
	4. Discussion/Conclusions
Conclusions	Statements of the specific conclusions, inferences, applications or generalizations drawn from the findings within a context of previous research and logic that clearly demonstrates and supports the researcher's derivation of these conclusion.
Limitations	The qualifications or limitations on the findings and conclusions that indicate clearly the parameters of the conclusions and to whom these conclusions apply and do not apply.
Alternative Hypotheses	A critical analysis of the research problem, design, procedures, cases and findings that indicates possible alternative interpretations of the findings.
Future Research	New, further or refined lines of research that are required to address the research problem more fully.

Figure 1-3. Expected general standards of high-quality educational research and questions to guide the evaluation and critique of research reports.

1. Specificity of The Research Problem

- Is the research problem precisely stated?
- Is the research problem properly integrated with theory and the research literature?
- Did the author reasonably delimit the problem?

2. Relevance

- Is the research problem important or significant?
- Did the author state the assumptions made to derive the problem?
- Is this study needed (i.e., can the problem be answered from previous research)?

3. Replicability

- Are the methods or procedures thoroughly presented?
- Are the definitions of terms reasonably or operationally defined?

4. Quantification or Specification of Data Obtained

- What are the data sources, and to what extend may generalizations be drawn from them?
- Are the variables precisely stated or are the data collected specifically defined?
- Are the variables valid?
- Are the variables reliably and validly "measured" or established on more than one source?

5. Objectivity

- Is the research/researcher objective or trustworthy?
- Are the design/procedures and data analysis thorough?

6. Scholarship

- Are the inferences and conclusions based upon intelligent interpretation of the data?
- Are **all** reasonable interpretations of the data discussed?
- Are the data and conclusions discussed in relation to the research literature?
- Does the researcher indicate significant research problems needing further research?