
EXPERIMENTAL RESEARCH IN THE FIELD EDUCATIONAL TECHNOLOGY

Dr. Suraksha Bansal
Department of Education
Gandhi Institute of Professional & Technical Studies
Meerut

Experimental methods have been used extensively for many years to conduct research in education and psychology. However, applications of experiments to investigate technology and other instructional innovations in higher education settings have been relatively limited or disfavor usage of experimental designs relative to other methods.

The experimental method formally surfaced in educational psychology around the turn of the century, with the classic studies by Thorndike and Woodworth on transfer. During the past century, the experimental method has remained immune to paradigm shifts in the psychology of learning, including behaviorism to cognitivism, objectivism to cognitivism, and instructive to constructivism. Clearly, the logical positivism of behaviorist theory created a fertile, inviting framework for attempts to establish causal relationships between variables, using experimental methods. The emergence of cognitive learning theory in the 1970s and 1980s initially did little to change this view, as researchers changed the locus of inquiry from behaviour to mental processing but maintained the experimental method as the basic way they searched for scientific truths. Today, the increasing influences of constructivist theories are making the fit between traditional scientific methods and current perspectives on learning more difficult, therefore, is to present experimental methods as continuing to provide valuable “tools” for research but ones whose uses may need to be altered or expanded relative to their traditional functions to accommodate the changing complexion of theory and scientific inquiry in instructional technology.

TYPES OF EXPERIMENTAL DESIGNS

In experimental research, the researcher manipulates or varies an independent variable and measures its effects on one or more dependent variables. In a true experimental design, the researcher randomly assigns the participants who are being studied (also called the subjects) to two or more comparison groups. Sometimes the comparison groups are referred to as treatment and control groups. Participants in the treatment group receive some type of treatment, such as a special reading program. Participants in the control group do not receive the treatment. Internal validity and external validity, and generally are not useful for making policy decisions.

ALTERNATIVE EXPERIMENTAL DESIGNS

They are also “core” designs in the sense of including basic components of the more complex or related designs not covered.

TRUE EXPERIMENTS

The ideal design for maximizing internal validity is the true experiment, as diagrammed below. The R means that subjects were randomly assigned, X represents the treatment (in this case, alternative

treatments 1 and 2), and O means observation (or outcome), for example, a dependent measure of learning or attitude. What distinguishes the true experiment from less powerful designs is the random assignment of subjects to treatments, thereby eliminating any systematic error that might be associated with using intact groups. The two (or more) groups are then subjected to identical environmental conditions, while being exposed to different treatments. In educational technology research, such treatments frequently consist of different instructional methods.

REPEATED MEASURES

A variation of the above experimental design is the situation where all treatments are administered to all subjects. Thus, each individual in essence, serves as his or her own control and is tested or “observed” as diagrammed below for an experiment using n subjects and k treatments. Note that the diagram shows each subject receiving the same sequence of treatments; a stronger design, where feasible, would involve randomly ordering the treatments to eliminate a sequence effect.

Suppose that an experimenter is interested in whether learners are more likely to remember words that are italicized or words that are underlined in a computer text presentation. Twenty subjects read a paragraph containing five words in each form. They are then asked to list as many italicized words and as many underlined words as they can remember. Note that this design has the advantage of using only one group, thereby effectively doubling the number of subjects per treatment relative to a two-group design. It also ensures that the ability level of subjects receiving the two treatments will be the same. But there is a possible disadvantage that may distort results. The observations are not independent. Recalling an italicized word may help or hinder the recall of an underlined word, or vice versa.

QUASI-EXPERIMENTAL DESIGNS

In a quasi-experimental design, the researcher does not randomly assign participants to comparison groups, usually because random assignment is not feasible. To improve a quasi-experimental design, the researcher can match the comparison groups on characteristics that relate to the dependent variable. For example, a researcher selects from a school district 10 classes to have low student-teacher ratios and 10 classes to maintain their current high student-teacher ratios. The researcher selects the high-ratio classes based on their similarity to the low-ratio classes in terms of student socioeconomic status, a variable that is related to student achievement.

QUASI- EXPERIMENTAL APPROACH IS TIME SERIES DESIGNS

This family of designs involves repeated measurement of a group, with the experimental treatment induced between two of the measures, thus a quasi-experiment is opposed to a true experiment. The absence of randomly composed, separate experimental and control groups makes it impossible to attribute changes in the dependent measure directly to the effects of the experimental treatment. That is, the individual group participating in the time series design may improve its performances from pretesting to post testing, but is it the treatment or some other event that produced the change. There is a variety of time series designs, some of which provide a higher internal validity than others.

DECEPTIVE APPEARANCES: THE EX POST FACTO DESIGN

Despite the appearances of a treatment comparison and random assignment, this research is not an experiment but rather an ex post facto study. No variables are manipulated. Existing groups that are essentially self-selected are being compare, those who chose the word processor vs. those who chose paper and pencil. The random selection merely reduced the number of possible participants to more manageable numbers; it did not assign students to particular treatments. Given these properties, the ex post facto study may look sometimes like an experiment but is closer in design to a co- relational study.

VALIDITY THREATS IN EXPERIMENTAL RESEARCH

By validity “threat,” we mean only that a factor has the potential to bias results. In 1963, Campbell and Stanley identified different classes of such threats.

- **Instrumentation.** Inconsistent use is made of testing instruments or testing conditions or the pre-test and post- test are uneven in difficulty, suggesting a gain or decline in performance that is not real.
- **Testing.** Exposure to a pre-test or intervening assessment influences performance on a post-test.
- **History.** This validity threat is present when events, other than the treatments, occurring during the experimental period can influence results.
- **Maturation.** During the experimental period, physical or psychological changes take place within the subjects.
- **Selection.** There is a systematic difference in subjects’ abilities or characteristics between the treatment groups being compared.
- **Diffusion of Treatments.** The implementation of a particular treatment influences subjects in the comparison treatment
- **Experimental Mortality.** The loss of subjects from one or more treatments during the period of the study may bias the results.

In many instances, validity threats cannot be avoided. The presence of a validity threat should not be taken to mean that experimental findings are inaccurate or misleading. Knowing about validity threats gives the experimenter a framework for evaluating the particular situation and making a judgment about its severity. Such knowledge may also permit actions to be taken to limit the influences of the validity threat in question

EXPERIMENTAL RESEARCH IN EDUCATIONAL TECHNOLOGY

Here is a sequence of logical steps for planning and conducting research

Step 1. Select a Topic- This step is self-explanatory and usually not a problem, except for those who are “required” to do research as opposed to initiating it on their own. The step simply involves identifying a general area that is of personal interest and then narrowing the focus to a researchable problem

Step 2. Identify the Research Problem- Given the general topic area, what specific problems are of interest? In many cases, the researcher already knows the problems. In others, a trip to the library to read background literature and examine previous studies is probably needed. A key concern is the importance of the problem to the field. Conducting research requires too much time and effort to be examining trivial questions that do not expand existing knowledge. Experienced researchers will usually be attuned to important topics, based on their knowledge of the literature and current research activities. Novices, however, need to be more careful about establishing support for their idea from recent research and issues-oriented publications.

Step 3. Conduct a Literature Search-

With the research topic and problem identified, it is now time to conduct a more intensive literature search. Of importance is determining what relevant studies have been performed; the designs, instruments, and procedures employed in those studies; and, most critically, the findings. Based on the review, direction will be provided for (a) how to extend or complement the existing literature base, (b) possible research orientations to use, and (c) specific research questions to address.

Step 4. State the Research Questions-

This step is probably the most critical part of the planning process. Once stated, the research questions or hypotheses provide the basis for planning all other parts of the study: design, materials, and data analysis. In particular, this step will guide the researcher’s decision as to whether an experimental design or some other orientation is the best choice.

Step 5. Determine the Research Design-

The next consideration is whether an experimental design is feasible. If not, the researcher will need to consider alternative approaches, recognizing that the original research question may not be answerable as a result.

Step 6. Determine Methods-

Methods of the study include (a) subjects, (b) materials and data collection instruments, and (c) procedures. In determining these components, the researcher must continually use the research questions and/or hypotheses as reference points. A good place to start is with subjects or participants. What kind and how many participants does the research design require?

Next consider materials and instrumentation. When the needed resources are not obvious, a good strategy is to construct a listing of data collection instruments needed to answer each question.

Step 7. Determine Data Analysis Techniques-

Whereas statistical analysis procedures vary widely in complexity, the appropriate options for a particular experiment will be defined by two factors: the research questions and the type of data

REPORTING AND PUBLISHING EXPERIMENTAL STUDIES

Obviously, for experimental studies to have impact on theory and practice in educational technology, their findings need to be disseminated to the field.

INTRODUCTION

The introduction to reports of experimental studies accomplishes several functions: (a) identifying the general area of the problem, (b) creating a rationale to learn more about the problem, (c) reviewing relevant literature, and (d) stating the specific purposes of the study. Hypotheses and/or research questions should directly follow from the preceding discussion and generally be stated explicitly, even though they may be obvious from the

LITERATURE REVIEW

In basic research experiments, usage of hypotheses is usually expected, as a theory or principle is typically being tested. In applied research experiments, hypotheses would be used where there is a logical or empirical basis for expecting a certain result

METHODOLOGY

The Method section of an experiment describes the participants or subjects, materials, and procedures. The usual convention is to start with subjects (or participants) by clearly describing the population concerned (e.g., age or grade level, background) and the sampling procedure. In reading about an experiment, it is extremely important to know if subjects were randomly assigned to treatments or if intact groups were employed. It is also important to know if participation was voluntary or required and whether the level of performance on the experimental task was consequential to the subjects. Learner motivation and task investment are critical in educational technology research, because such variables are likely to impact directly on subjects' usage of media attributes and instructional strategies

RESULTS

This major section describes the analyses and the findings. Typically, it should be organized such that the most important dependent measures are reported first. Tables and/or figures should be used judiciously to supplement (not repeat) the text. Statistical significance vs. practical importance traditionally, researchers followed the convention of determining the "importance" of findings based on statistical significance. Simply put, if the experimental group's mean of 85% on the post test was found to be significantly higher (say, at $p < .01$) than the control group's mean of 80%, then the "effect" was regarded as

having theoretical or practical value. If the result was not significant (i.e., the null hypothesis could not be rejected), the effect was dismissed as not reliable or important.

CRITERIA FOR REJECTION FOR PUBLICATION

Here are few reasons that makes an experimental study “publishable or perishable” in professional research journals.

Poor writing: Writing style is unclear, weak in quality (syntax, construction), and/or does not use appropriate (APA) style.

Invalid testing: Outcomes are not measured in a controlled and scientific way (e.g., observations are done by the author without validation of the system or reliability checks of the data).

Inappropriate analyses: Quantitative or qualitative analyses needed to address research objectives are not properly used or sufficiently described

Low internal validity of conditions: Treatment and comparison groups are not uniformly implemented. One or more groups have an advantage on particular condition (materials, and encouragement) other than the independent variable.

Low internal validity of subject selection/assignment: Groups assigned to treatment and comparison conditions are not comparable.

Low external validity: Application or importance of topic or findings is weak.

The research journals have published proportionately more experimental studies than alternative types. This factor also creates a self-perpetuating situation in which increased exposure to experimental studies increases the likelihood that beginning researchers will also favour the experimental method in their research.

CONTEMPORARY AREAS IN EDUCATIONAL RESEARCH EXPERIMENTATION

Randomized Field Experiments-

Given the importance of balancing external validity (application) and internal validity (control) in educational technology research, an especially appropriate design is the randomized field experiment, in which instructional programs are evaluated over relatively long periods of time under realistic conditions. In contrast to descriptive or quasi-experimental designs, the randomized field experiment requires random assignment of subjects to treatment groups, thus eliminating differential selection as a validity threat.

BASIC– APPLIED DESIGN REPLICATIONS

Basic research designs demand a high degree of control to provide valid tests of principles of instruction and learning. Once a principle has been thoroughly tested with consistent results, the natural progression is to evaluate its use in a real-world application. For educational technologists interested in how learners are affected by new technologies, the question of which route to take, basic vs. applied, may pose a real dilemma. Typically, existing theory and prior research on related interventions will be sufficient to raise the possibility that further basic research may not be necessary. To avoid the limitations of addressing one perspective only, a potentially advantageous approach is to look at both using a replication design. Consistency of findings across experiments would provide strong convergent evidence supporting the obtained effects and underlying theoretical principles. Inconsistency of findings, however, would suggest influences of intervening variables that alter the effects of the variables of interest when converted from their “pure” form to realistic applications. The next implied step of a replication design would be further experimentation on the nature and locus of the altered effects in the applied situation

ASSESSING MULTIPLE OUTCOMES IN EDUCATIONAL TECHNOLOGY EXPERIMENTS

In educational technology research, research questions are not likely to be resolved in straightforward a manner. Merely knowing that one instructional strategy produced better achievement than another provides little insight into how those effects occurred or about other possible effects of the strategies. Earlier

educational technology experiments, influenced by behaviouristic approaches to learning, were often subject to this limitation.

COGNITIVE APPLICATIONS

One key contribution has been the expansion of conventional assessment instruments so as to describe more fully the “cognitive character” of the target. Among the newer, cognitively derived measurement applications that are receiving greater usage in research are tests of declarative and procedural knowledge, componential analysis, computer simulations, faceted tests, and coaching methods, to name only a few. Whereas behavioural theory stressed learning products, such as accuracy and rate, cognitive approaches also emphasize learning processes. The underlying assumption is that learners may appear to reach similar destinations in terms of observable outcomes but take qualitatively different routes to arrive at those points. Importantly, the routes or “processes” used determine the durability and transferability of what is learned. Process measures may include such variables as the problem-solving approach employed, level of task interest, resources selected, learning strategies used, and responses made on the task. At the same time, the cognitive approach expands the measurement of products to include varied, multiple learning outcomes such as declarative knowledge, procedural knowledge, long-term retention, and transfer.

QUALITATIVE RESEARCH

In recent years, educational researchers have shown increasing interest in qualitative research approaches. Such research involves naturalistic inquiries using techniques such as in-depth interviews, direct observation, and document analysis. Presently, in educational technology research, experimentalists have been slow to incorporate qualitative measures as part of their overall research methodology. Item Responses vs. Aggregate Scores as Dependent Variables Consistent with the “expanded assessment” trend, educational technology experiments are likely to include dependent variables consisting of one or more achievement (learning) measures, attitude measures, or a combination of both types. In the typical case, the achievement or attitude measure will be a test comprised of multiple items. By summing item scores across items, a total or “aggregate” score is derived. To support the validity of this score, the experimenter may report the test’s internal-consistency reliability or some other reliability index. Internal consistency represents “equivalence reliability”- the extent to which parts of a test are equivalent. Depending on the situation, these procedures could prove limiting or even misleading with regard to answering the experimental research questions. A fundamental question to consider is whether the test is designed to measure a unitary construct or multiple constructs.

In the latter cases, internal consistency reliability might well be low, because students vary in how they perform or how they feel across the separate measures. Specifically, there may be no logical reason why good performances on, say, the “math facts” portion of the test should be highly correlated with those on the problem-solving portion.

It may even be the case that the treatments being investigated are geared to affect one type of performance or attitude more than another. Accordingly, one caution is that, where multiple constructs are being assessed by design, internal-consistency reliability may be a poor indicator of construct validity. More appropriate indexes would assess the degree to which (a) items within the separate subscales inter-correlate (subscale internal consistency), (b) the makeup of the instruments conforms with measurement objectives (content validity), (c) students answer particular questions in the same way on repeated administrations (test-retest reliability), and (d) subscale scores correlate with measures of similar constructs or identified criteria (construct or predictive validity). Separate from the test validation issue is the concern that aggregate scores may mask revealing patterns that occur across different subscales and items. We explore this issue further by examining some negative and positive examples from actual studies.

AGGREGATING ATTITUDE RESULTS

More commonly, educational technology experimenters commit comparable oversights in analyzing attitude data. When attitude questions concern different properties of the learning experience or instructional context, it may make little sense to compute a total score, unless there is an interest in an overall attitude score.

MEDIA STUDIES VS. MEDIA COMPARISONS

As confirmed by our analysis of trends in educational technology experimentation, a popular focus of the past was comparing different types of media-based instruction to one another or to teacher-based instruction to determine which approach was “best.” The fallacy or, at least, unreasonableness of this orientation, now known as “media comparison studies,”

For present purposes, these considerations present a strong case against experimentation that simply compares media. Specifically, two types of experimental designs seem particularly unproductive in this regard. One of these represents treatments as amorphous or “generic” media applications, such as CBI, interactive video and Web-based instruction. The focus of the experiment then becomes which medium “produces” the highest achievement. More recently, this type of study has been used to “prove” the effectiveness of distance education courses. A second type of inappropriate media comparison experiment is to create artificially comparable alternative media presentations, such that both variations contain identical attributes but use different modes of delivery.

Similarly, to learn about television’s “effects” as a medium, it seems to make more sense to use an actual television program, than a simulation done with a home video camera.

Deductive Approach: Testing Hypotheses about Media Differences. In this first approach, the purpose of the experiment is to test a priori hypotheses of differences between the two media presentations based directly on analyses of their different attributes

The rationale for these hypotheses would be based directly on analyses of the special capabilities (embedded attributes or instructional strategies) of each medium in relation to the type of material taught. Findings would be used to support or refute these assumptions.

Inductive Approach: Replicating Findings across Media. The second type of study, which we have called media replications, examines the consistency of effects of given instructional strategies delivered by alternative media. Consistent findings, if obtained, are treated as corroborative evidence to strengthen the theoretical understanding of the instructional variables in question as well as claims concerning the associated strategy’s effectiveness for learning. If inconsistent outcomes are obtained, methods and theoretical assumptions are re-examined and the target strategy subjected to further empirical tests using diverse learners and conditions. Key interests are why results were better or worse with a particular medium and how the strategy might be more powerfully represented by the alternative media. Subsequent developmental research might then explore ways of incorporating the suggested refinements in actual systems and evaluating those applications. In this manner, media replication experiments use an inductive, post hoc procedure to identify media attributes that differentially impact learning

For experimental studies to have an impact on theory and practice in educational technology, their findings need to be disseminated to other researchers and practitioners. Getting a research article published in a good journal requires careful attention to writing quality and style conventions. Typical write-ups of experiments include as major sections an introduction, method, results, and discussion.

REFERENCES

1. **Adaval, S.B., (1968)** -The Third Indian Year Book of Educational Research NCERT, Delhi, P-317.
2. **Agarwal, J.C. (2003)** -Research methodology, Vol.1 -2003.
3. **Ahmad, Q.(1986)**-“Determinants of job involvement among teachers” Ph.D Thesis on psychology, Magadh University.
4. **Alderman, D.E., (1979)**-“The effect of special preparation on scholastic aptitude test verbal score”.
5. **Alexander A., (1974)** -Schneider’s in the world book encyclopedia Vol.-12 Chicago.