An Example of Experimental Research

In the remainder of this chapter, we present a published example of experimental research. Along with a reprint of the actual study itself, we critique the study, identify its strengths, and discuss areas we think could be improved. We also do this at the end of Chapters 14, 19, 21, 22 and 23, in each case analyzing the type of study discussed in the chapter. In selecting the studies for review, we used the following criteria:

- The study had to exemplify typical, but not outstanding, methodology and permit constructive criticism.
- The study had to have enough interest value to hold the attention of students, even though specific professional interests may not be directly addressed.
- The study had to be concisely reported.

In total, these studies represent the diversity of special interests in the field of education.

In critiquing each of these studies, we used a series of categories and questions that should, by now, be familiar to you. They are:

Purpose/justification: Is it logical? Is it convincing? Is it sufficient? Do the authors show how the results of the study have important implications for theory, practice, or both? Are assumptions made explicit?

Definitions: Are major terms clearly defined? If not, are they clear in context?

Prior research: Has previous work on the topic been covered adequately? Is it clearly connected to the present study?

Hypotheses: Are they stated? Implied? Appropriate for the study?

Sample: What type of sample is used? Is it a random sample? If not, is it adequately described? Do the authors recommend or imply generalizing to a population? If so, is the target population clearly indicated? Are possible limits to generalizing discussed?

Instrumentation: Is it adequately described? Is evidence of adequate reliability presented? Is evidence of validity provided? How persuasive is the evidence or the argument for validity of inferences made from the instruments?

Procedures/internal validity: What threats are evident? Were they controlled? If not, were they discussed?

Data analysis: Are data summarized and reported appropriately? Are descriptive and inferential statistics (if any) used appropriately? Are the statistics interpreted correctly? Are limitations discussed?

Results: Are they clearly presented? Is the written summary consistent with the data reported?

Discussion/interpretations: Do the authors place the study in a broader context? Do they recognize limitations of the study, especially with regard to population and ecological generalizing of results?

How Manipulatives Affect the Mathematics Achievement of Students in Nigerian Schools

F. Ehi Aburime
Benson Idahosa University

Mathematics is a very important subject in Nigeria. Yet, for more than 20 years, mathematics education in Nigeria has been in a sorry state. Mathematics achievement has been very low and frustrating. So far, every effort made to save Nigerian education from the devastating effect of persistent poor mathematics achievement has failed. An experiment to address the problem of poor achievement in mathematics in Nigerian high schools was carried out in Edo State of Nigeria. Eighteen simple improvised geometric manipulatives were made from ordinary cardboard paper. The manipulatives were used in teaching students in experimental group. There was a control group of students which did not study with manipulatives. Scores were collected from mathematics test taken by students in both experimental and control groups. Statistical analysis showed that students in the experimental group (who were taught with manipulatives) were clearly better than students in the control group who were not instructed with manipulatives.

Too strongly worded

INTRODUCTION

Rationale and justification

Federal Republic of Nigeria (2004) National Policy on Education made mathematics compulsory in all classes in grade schools and high schools. In fact, in grade school and high school, every child must study mathematics everyday the child goes to school in Nigeria. The National Policy on Education also made it compulsory for students to pass mathematics at the end of junior high and the senior high school levels of education in order to continue their educational career. It is therefore necessary that mathematics be taught effectively in Nigerian schools. Another reason for desiring effective mathematics teaching in Nigerian schools is that mathematics is very much needed for undergraduate admission into universities in Nigeria (see Table 1 below).

TABLE 1 *Mathematics Admission Requirements for Nigerian Universities 2004–2005*

S/N	Faculty	Number of Courses Available	Courses Needing at Least Credit In Mathematics	Percent of Courses Needing at Least Credit in Mathematics
1	Administration	21	16	76%
2	Agriculture	46	46	100
3	Arts	65	0	0
4	Education	72	36	50
5	Engineering	68	68	100
6	Law	7	0	0
7	Medical Sciences	21	21	100
8	Science	79	79	100
9	Social Sciences	36	21	58
	Total	**415**	**287**	**69%**

Source: Joint Admissions and Matriculation Board (2004) Universities Matriculation Examination Brochure (2004–2005).

Table 1 shows that 69 percent (more than two-thirds) of the university courses available need mathematics as entry requirement. In the courses concerned, a candidate must pass mathematics at credit or distinction level, and also offer mathematics at the Universities Matriculation Examination (UME) before being considered for admission. Out of nine faculties, four faculties (Faculty of Agriculture, Faculty of Engineering, Faculty of Medical Sciences, and Faculty of Science) require credit or distinction level pass in mathematics for all courses before candidates can qualify for further screening for admission. In spite of this great need for high achievement in mathematics, mathematics achievement has remained very low for many years. Lassa (1981) pointed out the sorry state of mathematics education in Nigeria. Lassa's warning did not stop the continuous high failure in mathematics and its resulting frustration of students and embarrassment of teachers. Ale (1989) declared that the Mathematical Association of Nigeria was launching a War Against Poor Achievement in Mathematics (WAPAM), but WAPAM has not succeeded in solving the problem of poor achievement in mathematics in Nigerian high schools. Ale (2003), in his capacity as Director of the National Mathematical Center, Abuja, Nigeria, launched a Mathematics Improvement Program. Yet, the sad situation persists. To show how serious the situation is, Amoo (2001) brought out the following table (Table 2).

TABLE 2 *Performance in Mathematics in West African School Certificate Examination 1995–1997*

Year	Entry (%)	Candidates Who Sat for Exam	Credit or Distinction 1–6	Pass 7–8	Fail 9
1995	466,971	462,273 99.0	76080 16.5	185,931 40.2	200,262 43.3
1996	519,656	514,342 99.1	51587 10.0	190,899 37.1	272,356 52.9
1997	621,844	616,923 99.2	47252 7.7	161,526 26.2	408,145 66.2

Source: Amoo (2001).

In 1995, only 16.5 percent of high school students passed mathematics at the credit or distinction needed as a precondition for admitting students into majority of university courses. In 1996, the percent fell further to 10.0 percentage, only to fall again to 7.7 percent in 1997. Outright failure (Fail 9) in high school final year mathematics examination rose from 43.3 percent in 1995 through 52.9 percent in 1996 to 66.2 percent in 1997. This shows that nearly two-thirds of final year students in Nigerian high schools failed mathematics in 1997! These results are obviously not encouraging. They frustrate not only the students affected, but also other students. Mathematics teachers in Nigerian high schools as well as parents, guardians, and government are not happy about this persistent poor performance in mathematics. According to Ibuot (2000), a leading teacher, Okubodejo, said that:

> Government has not been happy with the performance of students in mathematics in recent times because of students' poor performance . . . Mathematics is the bedrock of the sciences and technology . . . Without mathematics it would be difficult for the nation to move forward. (p. 5)

A vivid description from Aborisade (2001) shows that,

> For many Nigerian students in secondary schools, mathematics is a loathsome subject... At the mention of the subject, some students curse and hiss. To those students who detest the subject, the mathematics teacher is an archenemy. He is

> unpopular simply because the subject is unpopular. A recently conducted research to know what is responsible for learners' hateful attitude to mathematics gave the reason to be lack of proper motivation and high cost of books. (p. 48)

Both Ale (1989, p. 27) and Amazigo (2000, p. 24) have identified teaching problems and lack of instructional materials as major factors responsible for poor performance in mathematics. As a contribution to the struggle against poor achievement in mathematics, an experiment was performed in Edo State, Nigeria, using three high schools—Momodu College, Agbede (mixed), Girls Model Secondary School, Ubiaja (girls only); Immaculate Conception College, Benin City (boys only). Simple improvised geometric manipulatives were utilized in the experiment. Manipulatives, according to National Science Foundation (2002), are materials designed to provide concrete experiences that can help students make the link between mathematical concepts and the real world. Also, Langa (2002) sees manipulatives as any objects that aid children in visualizing mathematical process. Ivowi (1999) explains that "the practical nature of science and mathematics is being emphasised in order to make them functional" (p. 483).

STATEMENT OF THE PROBLEM

In full recognition of the importance of mathematics, the Federal Government of Nigeria has made mathematics compulsory in all Nigerian grade sehools and high schools. Since government spends a lot of money on education, it is disappointing to see a high percentage of students continuously failing mathematics for more than 20 years. In spite of efforts made by the Mathematical Association of Nigeria and the National Mathematical Centre, Abuja, Nigeria, the problem has persisted. Improving achievement in mathematics is desirable. As a matter of fact, it is necessary in order to produce more students who can acquire enough mathematics education to carry Nigeria confidently to the forefront of scientific and technological research and development in this twenty-first century. As a contribution toward solving the longstanding problem of low achievement in mathematics in Nigerian high schools, an experimental research study was carried out in Edo State, Nigeria. The experiment aimed to test the effect of 18 simple geometric manipulatives, made from ordinaiy cardboard paper, on the mathematics achievement of students in Edo State.

OBJECTIVE OF THE STUDY

Clear purpose

The study aims to find out how simple improvised geometric manipulatives affect mathematics achievement of high school students. Specifically, the study wants to determine if mathematics achievement in Edo State high schools in Nigeria will improve as a result of using improvised geometric manipulatives in classroom teaching and learning.

SIGNIFICANCE OF THE STUDY

Anticipated use of results

Once it is ascertained that improvised geometric manipulatives improve mathematics achievement at the high school level, teachers and students will be encouraged to use improvised manipulatives as one possible means of fighting failure and frustration in mathematics teaching and learning. There is a more important reason for encouraging the use of improvised geometric manipulatives (if they are effective enough to improve mathematics achievement). If more students pass mathematics, mathematics can become a more solid springboard for economic, scientific, and technological development.

RESEARCH QUESTION

To what extent does the use of improvised geometric manipulatives improve the achievement of students in mathematics?

Research Hypothesis

There is no significant difference between the mathematics achievement of students taught with simple improvised geometric manipulatives and students taught without manipulatives.

Could be expected outcome, not null hypothesis

Research Methodology

Population, sample, instrument, data collection and analysis shall be discussed under research methodology.

Population

All students in 287 high schools in Edo State of Nigeria constitute the population for this study.

Sample

Stratified random sampling, using a table of random numbers, led to the choice of Momodu College, Agbede (a mixed school), Girls Model Secondary School, Ubiaja (girls only), and Immaculate Conception College, Benin City (boys only). A random choice of class was done by simple ballot with JSSI, JSS2, JSS3, SSS1, SSS2, SSS3 written, one each, on six pieces of paper of equal size. After folding the papers and throwing them on a wide floor, one of them was picked up. It was JSS2. Another simple ballot was conducted with two equal pieces of paper, with A written on one paper, and B written on the other paper. The papers were folded and thrown into an envelope. After closing and shaking the envelope, one of them was taken out. Before picking the paper, it was decided that the chosen paper should be for the experimental group while the remaining paper would be for the control group. B was the choice for experimental group. So JSS2B became the experimental group while JSS2A was the control group in all the three schools chosen.

Not clear to us

Randomizes classes, not students

Instrument

The instrument for this research consisted of a testing instrument, Mathematics Achievement Test (MAT) and a treatment instrument, Simple Improvised Manipulatives (SIM). MAT was a mutiple-choice type of achievement test, which was used for both pretest and posttest. To construct MAT, a table of specifications (or test blueprint) was drawn up for 68 test items (see Table 3 below).

TABLE 3 *Specifications for JSS2 Mathematics Achievement Test*

Mathematics Areas	Recall of Information	Understanding Concepts	Applications of Concepts	Total
Geometry	5	13	8	26
Algebra	3	7	4	14
Statistics	3	7	4	14
Number and Numeration	3	7	4	14
Total	**14**	**34**	**20**	**68**

Adds clarity

Writing of test items was followed by face and content validation, then item analysis. The face and content validation reduced the items from 68 to 61 while item analysis reduced the test items from 61 to 52. The 52 surviving items were administered to 31 students in a pilot test. The students were in JSS2 in Obiaruku Grammar School, Obiaruku, Nigeria. Kuder-Richardson Formula (KR21) was applied to scores in order to measure internal consistency. The internal consistency coefficient was 0.79. It was considered high enough to accept MAT for research.

By whom and how?

Good

Adequate reliability

A test-retest of students, with an interval of three weeks, yielded scores that were paired and analyzed to obtain 0.83 as test-retest reliability coefficient for MAT. Again, this was high enough to accept MAT as a reliable research instrument. The second type of research instrument was made up of SIM. SIM had 18 different geometrical shapes constructed from ordinary cardboard paper. Four shapes were triangles—equilateral triangle, isosceles triangle, right-angled triangle, and scalene triangle. Six shapes were quadrilaterals—square, rhombus, rectangle, parallelogram, trapezium, and kite. There were four other plane shapes—pentagon, hexagon, circle, and semicircle. Cube, cuboid, triangular prism, and cylinder were the four solid shapes in SIM.

What was done with these?

DATA COLLECTION AND ANALYSIS TECHNIQUES

Need more description

Internal validity

Good

Data collection started with a pretest administered to students in both experimental group and control group. The pretest scores were carefully kept for future use. After pretest, 10 weeks of teaching followed, during which only the experimental group was treated with SIM. In other words, during the 10 weeks that followed the pretest, students in the experimental group were taught with simple improvised manipulatives while control group students were taught without using manipulatives. Special care was taken to make sure that, in each school, the same mathematics teacher taught the experimental group (JSS2B) and the control group (JSS2A). This was to eliminate what is known as the “teacher effect.” Another teaching precaution was to make sure that each time one group (JSS2A or JSS2B) was having a mathematics class, the other group was occupied by another teacher. For example, when JSS2A was studying mathematics, JSS2B must be studying mathematics under a teacher, so that JSS2B students could not study mathematics in JSS2A classroom. In the same way, whenever JSS2B was having a mathematics lesson, a teacher kept JSS2A occupied with another subject. These two precautions were in operation throughout the 10 weeks of teaching. After the pretest and 10 weeks of teaching, a posttest was given to all students in experimental group as well as control group. The posttest scores were collected for analysis in conjunction with pretest scores. It is noteworthy that the validated MAT was used for both pretest and posttest with several precautions. The first precaution was that MAT was administered to students in two versions. The second version of MAT was a rearrangement of the numbers and alternative answers in the first version. Thus, question number 5 with correct answer E in the first version of MAT could become question number 23 with correct answer C in the second version of MAT. The second precaution was to minimize the chances of obtaining fake scores from students who merely copy from their neighbors. To carry out this precaution, question papers were given to students in a checkerboard fashion (see Table 4 below).

Checkerboard arrangement of question papers guarantees that students writing a different version of test surround every student. Also, desks were separated far enough to prevent any student copying diagonally from students writing the same version. The third precaution

TABLE 4 *Checkerboard Arrangement of Question Papers*

1	2	1	2	1	2	1	2	1	2
2	1	2	1	2	1	2	1	2	1
1	2	1	2	1	2	1	2	1	2
2	1	2	1	2	1	2	1	2	1
1	2	1	2	1	2	1	2	1	2
2	1	2	1	2	1	2	1	2	1

1. First Version Test 2. Second Version Test

was that no student wrote the same version of MAT for pretest and posttest. Statistical analysis of pretest scores was carried out by the computer, using *t*-test (see tables below).

FINDINGS

Table 5A shows the number of students in experimental group (JSS2B) and control group (JSS2A) in the three schools where the research took place. The experimental group had a total of 94 students while the control group had 91 students altogether.

TABLE 5A *Number of Students in Experimental and Control Groups*

	Immaculate Conception College, Benin City	Girls Model High School, Ubiaja	Momodu College, Agbede	Total
Group JSS2B	36	36	22	**94**
Group JSS2A	33	35	23	**91**
Total	**69**	**71**	**45**	**185**

Table 5B shows that, at the pretest level, the calculated *t*-value (.09) is very much below the critical t value (1.96) at .05 level of significance. Therefore, there is no significant difference between the mean score (8.74) of the experimental group and the mean score (8.81) of the control group. This implies that the experimental and control groups were academically equal in mathematics achievement at the pretest level.

Does not establish equality

At the posttest level, the calculated *t*-value (2.23) is greater than the critical t value (1.96) as shown in Table 5C. This implies that a significant difference exists between the mean (11.70) of the experimental group and the mean (9.89) of the control group. The experimental group is now clearly superior to the control group in mathematics achievement at the posttest level. Thus, the research hypothesis (claiming no significant difference) has to be rejected in favor of the experimental group with much higher posttest mean score than the control group. This goes to show that students taught with simple improvised geometric manipulatives performed much better than other students.

Too strong

Effect size is needed (Delta)

TABLE 5B t-*test Analysis of Pretest Scores of Experimental and Control Groups*

Group	Number	Mean	Standard Deviation	df	Calculated *t*-value	Critical *t*-value
Experimental (JSS2B)	94	8.74	5.37	184	0.9*	1.96
Control (JSS2A)	91	8.81	5.18			
Decision on Hypothesis			**Accept Hypothesis**			

*Not significant at .05 level.

TABLE 5C *t-test Analysis of Posttest Scores of Experimental and Control Groups*

Group	Number	Mean	Standard Deviation	df	Calculated *t*-value	Critical *t*-value
Experimental (JSS2B)	94	11.70	5.26	184	2.23*	1.96
Control (JSS2A)	91	9.89	5.77			
Decision on Hypothesis			**Reject Hypothesis**			

*significant at .05 level.

DEFINITION OF KEY TERMS

Control Group	Students who were taught without using manipulatives
Credit	B or C grade
df	Degrees of freedom
Distinction	A grade
Experimental Group	Students who were taught with simple improvised geometric manipulatives
Item	Question
JSS	Junior secondary school (junior high school)
Manipulatives	Instructional materials that learners can easily handle to help them visualize and understand mathematical ideas
Secondary School	High school
SSS	Senior secondary school (senior high school)

RECOMMENDATIONS

Independent of this study

1. Nigerian government should make it a point of duty to arrest the present persistent poor performance in high school mathematics. Poor performance in high school mathematics has disturbed Nigeria for many years. It is inimical to national development.
2. The use of simple, improvised manipulatives for mathematics teaching and learning should be introduced in our schools, and entrenched in our curriculum. This should assist educational development in particular, and national development in general.
3. To sustain democracy and develop this country to a high level, our federal, state, and local govemments should invest massively in education, especially mathematics education which is the basis for development in science and technology.

CONCLUSION

Too strong

The result of this research study has shown that the persistent poor performance of our high school students in mathematics for many years need not continue indefinitely. There is hope that, with simple, cheap, improvised manipulatives, the situation can be changed for the better. With ordinary cardboard paper, teachers and students can construct and use simple geometric manipulatives to improve mathematics teaching and learning. Better mathematics results will follow; and better-trained students will come out of Nigerian high schools. Such students will be more equipped for national development, especially in the fields of science and technology

Too strongly worded

where mathematics is a necessity, and not a luxury. Moreover, better mathematics education due to improvement in mathematics achievement will raise the educational quality of Nigeria and lead the country to sustainable development. In this age of the computer, the website, and the Internet, education needs to improve so that Nigeria will not be left behind by other nations in this highly competitive and dynamic world.

References

Aborisade, A. (2001, February 19). *The wonder of mathematics*. Tell.

Ale, S. O. (1989). Combating poor achievement in mathematics (A keynote speech). *Abacus, 1*(1), 26–40.

Ale, S. O. (2003, January 29). Global view of Mathematics Improvement Program (MIP) and its mission. (A welcome address). Opening ceremony, orientation programme for mathematics improvement, National Mathematical Center, Abuja, Nigeria.

Amazigo, J. C. (2000, July 12). Mathematics phobia: Diagnosis and prescription. First Annual Lecture, National Mathematical Center, Abuja, Nigeria.

Amoo, S. A. (2001, September 20). Curriculum ideals and realities for sustainable educational development: A case of high school mathematics education curriculum. Paper presented at the 14th Annual National Conference of the Curriculum Organization of Nigeria, Abuja, Nigeria.

Federal Republic of Nigeria. (2004). *National policy on education* (4th ed.). Lagos: NERDC Press.

Ibuot, U. (2000, July 9). Mathematics teachers seek government support. *Sunday Vanguard*, 5.

Ivowi, U. M. O. (1999). The philosophy and objectives of the science and mathematics curricula at the SS level. In B. B. Akpan, (Ed.), *Perspectives on education and science teaching: From the eyes of Uduogie Ivowi*. Ibadan: Science Teachers Association of Nigeria.

Joint Admissions and Matriculation Board. (2004). Universities Matriculation Examination/Direct Entry Brochure: Guidelines for Admission for First Degree Courses in Nigerian Universities, 2004–2005 Session. Lagos: Joint Admissions and Matriculation Board.

Langa, M. A. (2002). *Manipulatives: An exciting tool for learning math*. Retrieved October 22, 2005, from www.parenthhood.com/library/langa_mathfun.htm.

Lassa, P. N. (1981, January 20). The sorry state of mathematics education in Nigeria. An address delivered at University of Jos, Nigeria.

National Science Foundation. (2002). *Elementary and secondary education curriculum and instruction*. Arlington, VA: Science and Engineering Indicators.

Analysis of the Study

PURPOSE/JUSTIFICATION

The purpose is clearly stated under "Objective of the Study": to determine whether mathematics achievement in Edo State high schools would improve by using improvised geometric manipulations. An extensive justification is given in the Introduction and Significance sections. The justification relies heavily on data showing poor achievement and on authorities recommending use of manipulatives.

PRIOR RESEARCH

No research on manipulatives is cited, which is a weakness in this published study. An extensive body of empirical research supports this method and would have provided additional justification for this study.

DEFINITIONS

Definitions of some Key Terms is provided but not for "mathematics achievement" or "geometric manipulates." However, we think the former is made clear in the "operational" description of MAT (see p. 31 of study). The second is more serious—although the materials themselves, that is, SIM, are well described, how they were used is not described here or elsewhere.

HYPOTHESES

The research hypothesis is clearly stated although we think it is preferable to state it in terms of expected outcome rather than as the statistical "null" hypothesis (for more, see Chapter 11 in this textbook).

SAMPLE

The sample selection is unclear to us. It appears the schools were not chosen at random, which would have been a poor method regardless as selecting a random sample of three from a population of 287 is insufficient to produce representativeness. Whatever selection method was used should have been described. Randomness was properly used in assigning classes to experimental and control groups, one each at each school but it is not clear that these were the only classes. The report should have acknowledged the serious limitations to generalizing from the actual sample of 185 students to any population including the one stated—all high schools in Edo State.

INSTRUMENTATION

The format and development of the MAT are well described. The report should have stated who did the face and content analyses. Both internal consistency and re-test values are, as stated, acceptable for research purposes. Validity was enhanced by minimizing cheating and by changing the format from pre- to posttest.

INTERNAL VALIDITY

Although the study is a true experiment in that the sample was randomly divided into experimental/treatment and control groups (hereafter E and C), the desired equating of groups is greatly reduced because it was classes, not individuals, which were randomly divided. The group at each school could ideally have been so divided but this may not have been feasible. Nevertheless, division of groups into E and C did result in adequate equating on the most important variable, prior math ability, as shown by MAT scores. It is, however, the similarity in means and standard deviations that matters here; inference statistics are intended to assess generalizability, and not to equate groups. A nonsignificant difference does not establish equality (see Chapter 11).

The means of selecting schools was apparently successful in equating E and C groups on gender though we can't be sure that the Momodu school was equally divided. Two other important variables, both procedural, were well controlled—teacher effect and contamination of E and C groups.

DATA ANALYSIS AND RESULTS

Data analysis consisted of *t*-tests calculated separately for E and C groups. Significance tests address generalizability and require random selection from defined groups, which is not the case in this study. They should not be used to assess importance of change; this should be done with Delta (see Chapter 12 in this textbook). The E group gain of 2.96 results in a delta of .56, high enough to be considered important. However, this must be considered along with the C group gain of 1.08, which is the reason for a control group. The Delta for gain was not obtained and cannot be calculated by a reader because the standard deviation of gain scores is not given (see Chapter 12). A tentative assessment can be made by comparing the difference in gain scores to the reported S.D.s which gives Deltas of .33 to .36—below the usual standard of .50. However, the S.D. of gain scores can be expected to be considerably smaller which would result in larger Deltas. It is also worth noting that the E group gain is three times that of C. We conclude that the data support the value of the treatment, but not as strongly as the author states.

RECOMMENDATIONS/CONCLUSIONS

Recommendations 1 and 3 are persuasive but are independent of this study. We think the conclusions are too strong and fail to acknowledge the several limitations of the study. We also think the data justify replication of the study to further assess effectiveness and generalizability to the intended population.

Go back to the **INTERACTIVE AND APPLIED LEARNING** feature at the beginning of the chapter for a listing of interactive and applied activities. Go to **McGraw Hill Connect®** to take quizzes, practice with key terms, and review chapter content.

Main Points

THE UNIQUENESS OF EXPERIMENTAL RESEARCH

- Experimental research is unique in that it is the only type of research that directly attempts to influence a particular variable, and it is the only type that, when used properly, can really test hypotheses about cause-and-effect relationships. Experimental designs are some of the strongest available for educational researchers to use in determining cause and effect.

ESSENTIAL CHARACTERISTICS OF EXPERIMENTAL RESEARCH

- Experiments differ from other types of research in two basic ways—comparison of treatments and the direct manipulation of one or more independent variables by the researcher.

RANDOMIZATION

- Random assignment is an important ingredient in the best kinds of experiments. It means that every individual who is participating in the experiment has an equal chance of being assigned to any of the experimental or control conditions that are being compared.

CONTROL OF EXTRANEOUS VARIABLES

- The researcher in an experimental study has an opportunity to exercise far more control than in most other forms of research.
- Some of the most common ways to control for the possibility of differential subject characteristics (in the various groups being compared) are randomization, holding certain variables constant, building the variable into the design, matching, using subjects as their own controls, and using analysis of the covariance.

POOR EXPERIMENTAL DESIGNS

- Three weak designs that are occasionally used in experimental research are the one-shot case study design, the one-group pretest-posttest design, and the static-group comparison design. They are considered weak because they do not have built-in controls for threats to internal validity.
- In a one-shot case study, a single group is exposed to a treatment or event, and its effects are assessed.
- In the one-group pretest-posttest design, a single group is measured or observed both before and after exposure to a treatment.
- In the static-group comparison design, two intact groups receive different treatments.

TRUE EXPERIMENTAL DESIGNS

- The essential ingredient of a true experiment is random assignment of subjects to treatment groups.
- The randomized posttest-only control group design involves two groups formed by random assignment.
- The randomized pretest-posttest control group design differs from the randomized posttest-only control group only in the use of a pretest.

- The randomized Solomon four-group design involves random assignment of subjects to four groups, with two being pretested and two not.

MATCHING

- To increase the likelihood that groups of subjects will be equivalent, pairs of subjects may be matched on certain variables. The members of the matched groups are then assigned to the experimental and control groups.
- Matching may be either mechanical or statistical.
- Mechanical matching is a process of pairing two persons whose scores on a particular variable are similar.
- Two difficulties with mechanical matching are that it is very difficult to match on more than two or three variables, and that in order to match, some subjects must be eliminated from the study when no matches can be found.
- Statistical matching does not necessitate a loss of subjects.

QUASI-EXPERIMENTAL DESIGNS

- The matching-only design differs from random assignment with matching only in that random assignment is not used.
- In a counterbalanced design, all groups are exposed to all treatments, but in a different order.
- A time-series design involves repeated measurements or observations over time, both before and after treatment.

FACTORIAL DESIGNS

- Factorial designs extend the number of relationships that may be examined in an experimental study.

Key Terms

comparison group 260
control 262
control group 260
counterbalanced design 269
criterion variable 259
dependent variable 259
design 262
experiment 260
experimental group 260
experimental research 259
experimental variable 259
extraneous variable 261
factorial design 271
gain score 269
independent variable 259
interaction 271
matching design 266
matching-only design 269
mechanical matching 268
moderator variable 271
nonequivalent control group design 264
one-group pretest-posttest design 263
one-shot case study design 263
outcome variable 259
pretest treatment interaction 266
quasi-experimental design 269
random assignment 261
random selection 261
randomized posttest-only control group design 265
randomized pretest-posttest control group design 265
randomized Solomon four-group design 266
regressed gain score 269
static-group comparison design 264
static-group pretest-posttest design 264
statistical equating 268
statistical matching 268
time-series design 270
treatment variable 259

For Discussion

1. An occasional criticism of experimental research is that it is very difficult to conduct in schools. Would you agree? Why or why not?
2. Are there any cause-and-effect statements you can make that you believe would be true in most schools? Would you say, for example, that a sympathetic teacher "causes" elementary school students to like school more?
3. Are there any advantages to having more than one independent variable in an experimental design? If so, what are they? What about more than one dependent variable?
4. What designs could be used in each of the following studies? (*Note:* More than one design is possible in each instance.)
   a. A comparison of two different ways of teaching spelling to first-graders.
   b. An assessment of the effectiveness of weekly tutoring sessions on the reading ability of third-graders.
   c. A comparison of a third-period high school English class taught by the discussion method with a third-period (same high school) English class taught by the lecture method.
   d. The effectiveness of reinforcement in decreasing stuttering in students with this speech defect.
   e. The effects of a year-long weight-training program on a group of high school athletes.
   f. The possible effects of age, gender, and method on student liking for history.
5. What flaw can you find in each of the following studies?
   a. A teacher tries out a new mathematics textbook with her class for a semester. At the end of the semester, she reports that the class's interest in mathematics is markedly higher than she has ever seen it in the past with other classes using another text.
   b. A teacher divides his class into two subgroups, with each subgroup being taught spelling by a different method. Each group listens to the teacher instruct the other group while they wait their turn.
   c. A researcher calls for eighth-grade students to volunteer to tutor third-grade students who are having difficulty in reading. She compares their effectiveness as tutors with a control group of students who are assigned to be tutors (they do not volunteer). The students of the volunteers have a much greater level of improvement in reading than the students who were assigned to tutor.
   d. A teacher decides to try out a new textbook in one of her social studies classes. She uses it for four weeks and then compares this class's scores on a unit test with the scores of her previous classes. All classes are studying the same material. During the unit test, however, a fire drill occurs, and the class loses about 10 minutes of the time allotted for the test.
   e. Two groups of third-graders are compared with regard to running ability, subsequent to different training schedules. One group is tested during physical education class in the school gymnasium, while the other is tested after school on the football field.
   f. A researcher compares a third-period English class with a fifth-period chemistry class in terms of student interest in the subject taught. The English class is taught by the discussion method, while the chemistry class is taught by the lecture method.

References

1. Nye, B., et al. (2001). Are effects of small classes cumulative? Evidence from a Tennessee experiment. *American Educational Research Journal, 37*(1), 123–151.

2. Lo, Y., et al. (2009). Examining the impacts of early reading intervention on the growth rates in basic literacy skills of at-risk urban kindergarteners. *Journal of Special Education, 43,* 12–28.

3. Stanulis, R. N., & Floden, R. E. (2009). Intensive mentoring as a way to help beginning teachers develop balanced instruction. *Journal of Teacher Education, 60*(3), 112–122.

4. Heerwegh, D. (2006). An investigation of the effect of lotteries on Web survey response rates. *Field Methods, 18*(5), 205–220.

5. Benitez, J. L. et al. (2009). The impact of a course on bullying within the pre-service teacher training curriculum. *Electronic Journal of Research in Educational Psychology, 7*(1), 191–207.

6. Kalyva, E., & Agaliotis, I. (2009). Can social stories enhance the interpersonal conflict resolution skills of children with LD? *Research in Developmental Disabilities: A Multidisciplinary Journal, 30*(7), 192–202.

7. DeVos, H. M., & Louw, D. A. (2009). Hypnosis-induced mental training programmes as a strategy to improve the self-concept of students. *Higher Education: The International Journal of Higher Education and Educational Planning, 57*(2), 141–154.

8. Tuckman, B. W. (1999). *Conducting Educational Research* (5th ed., p. 152). New York: College Publishers.