DOCUMENT RESUME

ED 451 206 TM 032 433

AUTHOR	Inoue, Yukiko
TITLE	Research Design and Analysis: Examples of Questions and Answers.
PUB DATE	2001-00-00
NOTE	9p.
PUB TYPE	Reports - Descriptive (141)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Data Analysis; *Research Design; Research Methodology; *Test Construction

ABSTRACT

This document provides some examples of questions that might be asked about research design and analysis, or that might be used in test construction. The accompanying answers form a basic discussion of research design and analysis techniques. Short-answer constructed response answers are provided for questions about: (1) control in an experiment; (2) statistical conclusion validity; (3) internal and external validity; (4) analysis of variance (ANOVA); and (5) ANOVA with one dependent variable and three independent variables. An essay question is posed and answered for a scenario involving teacher evaluation. An example problem related to ANOVA is also presented. (SLD)

Research Design and Analysis: Examples of Questions and Answers

Yukiko Inoue, PhD
Department of Foundations and Educational Research
College of Education
University of Guam
UOG Station, Mangilao, Guam 96923
Phone: 671-735-2447 Fax: 671-734-3651
Email: yinoue@uog.edu

Research Design and Analysis: Examples of Questions and Answers

1 SHORT ANSWER

1-1 What are the purposes of control in an experiment?

The basic purpose of control in an experiment is to try to control as many independent variables as possible in order to compare treatment groups against non-treatment groups. All of the various forms of relevant extraneous variables must be controlled. If they are not, the results of the experiment will be uninterpretable. This means that the internal validity of the experiment has been lost, and experiments with confounded results are of little value. The procedures that investigators use to control extraneous variables include: Eliminating variables, holding variables constant, balancing, systematic variation of variables, setting the extraneous variable against the independent variable, randomization, matching, and control of experimenter effect.

1-2 Define "statistical conclusion validity" and then list and explain the threats to it.

Statistical conclusion validity is defined as the validity of conclusions we draw on the basis of statistical evidence about whether a presumed cause and effect co-vary. Applying a hypothesis test may lead to the wrong conclusion. There are, in fact, two kinds of error possible called Type I error (rejecting H_o when H_{o} is true) and Type II error (not rejecting H_o when H_o is false). 1 - ß is referred to as the "power" of the statistical test, and power is the probability of rejecting the wrong presumption.

There are many ways to increase power and some of major considerations and threats can be explained as follow.

- Sample size: Studies have low power because sample size used is too small for situation.
- The alpha level: Studies have low power because significant level is too small for the sample size; for example, they used .01 instead of .05.
- Reliable dependent variable: Studies have low power because they used an unreliable dependent variable (reliability has to do with consistency and accuracy). For example, increasing the length of test increase their reliability.
- Random error: Unexplained variability in the dependent variable may be unacceptably high (e.g., implementing the treatment in different ways from one subject to the next).
- Statistical assumptions: Violations of statistical assumptions can also affect Type I and Type II error rates.

1-3 Discuss "internal validity" and "external validity" and then list and explain the threats to each.

Problems of internal validity are amenable to solution through the careful design of experiments, but this is not as true for external validity. External validity is largely a matter of generalization; thus, this is an inductive process of extrapolating beyond the data collected. In sum, internal validity refers to the validity of any conclusions we draw about whether a demonstrated statistical relationship implies cause, whereas external validity refers to the validity with which a causal relationship can be generalized across persons, settings, and times. Major threats to each are listed and explained below.

Threats to internal validity

- History: This refers to extraneous incidents or events affecting the results that occur during the research.
- Maturation: This refers to change in the subjects of a study over time such as getting hungry, tired, or discouraged.
- Testing: One form of this is the learning effect by which people improve on taking a second test even if it is an alternative form of the original.
- Instrumentation: This results from changes, between observations, in the measuring instruments, or observers.
- Selection: Systematic differences in selection of subjects.
- Statistical regression: This refers to the tendency of subjects who score very high or low on a pretest to scores close to the mean on the posttest.
- Mortality: Attrition is likely in the experimental group; each dropout changes the makeup of the group.

Threats to external validity

- Interaction of treatment and treatment: This occurs if respondents experience more than one treatment.
- Interaction of testing and treatment: In an experimental pretest we may sensitize subjects so that they respond to the experimental stimulus in a different way.
- Interaction of selection and treatment: This is a question of generalizing to other categories of people beyond the groups upon which the original relationship was founded.
- Interaction of setting and treatment: Unwillingness of some organizations to participate in a study may also promote the use of settings, which are different from the average.
- Interaction of history and treatment: Sometimes major events, which occur during the study, have the potential to confound treatment effects.

1-4 Discuss the considerations for deciding when to use a one-way ANOVA with random assignment to groups or analysis of covariance (one covariate and one independent variable).

A one-way analysis of variance (ANOVA) is a procedure to test the hypothesis that several populations have the same mean. The null hypothesis for a one-way ANOVA is H_o: $\mu_1 = \mu_2 = \mu_3 = \ldots = \mu_k$, and the alternative hypothesis is H_a: one or more of the population means is not equal to the others. A one-way ANOVA, in effect, is an extension of *t* Test and is to compare groups in terms of the mean scores. However, for example, in the case that students in grades 7, 8, 9, 10, 11, and 12 are compared on absenteeism, if ANOVA were used rather than multiple *t* Tests, the probability of Type I error would be less. Additionally, ANOVA technique makes certain assumptions about the data being investigated. The three major assumptions for the ANOVA are normality, homogeneity of variance, and independence of errors.

In analysis of covariance (ANCOVA), on the other hand, we combine the basic idea of analysis of variance and correctional analysis. Including a covariate does affect the analysis in two ways: 1) the within-group variability will be reduced by an amount dependent on the strength of the relationship between the depended variable and the covariate, and 2) the adjustment of the estimated magnitude of the treatment effect itself. For instance, in the case of studying the effect of SAT on students performance (freshman in college), including IQ as a covariate is a method of controlling the potential confounding effect of initial group ability differences. However, although ANCOVA can be used in an effort to make more nearly comparable intact groups that differ in known ways, always we have to remember that the adjustment may well introduce or exaggerate differences along some dimensions while it reduces the differences along other dimensions.

1-5 *Explain why in a research study with one dependent variable and three independent variables it is advantageous to use a single three factor ANOVA rather than a series of one or two factor ANOVAs.*

Considering that factorial designs introduce the concept of interaction, with the addition of a third factor, we can generalize the concept of an interaction because it may happen that all three factors interact. For example, in a study that we want to study the effects of three variables (the gender of the subject, the amount of prior experience, and the amount of marijuana) on rotary pursuit performance, the effects of these three variables and their interactions could be evaluated in a three-factor experiment. This procedure provides information about three main effects, three two-way interactions, and one three-way interaction as described below (and thus we have seven hypotheses).

Main effect for variable A (gender)
Main effect for variable B (prior experience)
Main effect for variable C (amount of marijuana)
A x B interaction
A x C interaction
B x C interaction
A x B x C interaction

Each of the two-way interactions is obtained by considering two variables at a time and averaging the values for the third variable. The three-way interaction, on the other hand, is obtained by considering all three variables at a time; this indicates whether the effect one variable has on the dependent variable is influenced by the combined presence of the other tow variables. Quite often knowledge concerning the three-way interaction helps to clarify the complex relationship among variables and provides important insight concerning the effect of each variable. This is a unique advantage of three-factor ANOVA. Of course, two-factor experiments are usually easier to interpret than three-factor experiments.

2 ESSAY

2-1 Critique (positive and negative) the following research design for this scenario. Your discussion should included validity (internal, external, statistical conclusion, and construct), control and causation and the use of randomization.

Scenario:

Professor Morgan is serving the Guam school system as a consultant on a beginning teacher evaluation program. There are two different approaches to evaluating teacher effectiveness. Both are based on a professionally accepted list of 12 teacher competencies and both used trained observers of classroom activity. Both evaluation methodologies require that two classes be observed on one day and there are three different observation days. Method 1 is referred to as "scripted" methodology in which the observer takes notes for a class period and afterward ratings of 1 - 4 are assigned (1=Unsatisfactorily, 2=Needs Improvement, 3=Area of Strength, and 4=Demonstrates Excellence). Method 2 is referred to as "sign" methodology in which about 250 very specific behaviors are "bubbled" if observed (otherwise left blank). From these behaviors about 40 raw scores/ratings are converted into a single composite scale from 20 to 80 which is then used to classify a teacher as I teaching license denied, II probationary license granted for two years, or III competent and full teaching license granted for ten years.

Research Question:	Do these two approaches to beginning teacher evaluation yield equivalent results?
Secondary Question:	Are these methods equivalent for difference subgroups? For example: Racial groups Genders School level (elementary, middle, high) Urban, suburban, rural

To address these questions a pilot study was planned. To ensure representativeness of the Guam school system, selection of teachers was to be done within each of the five regions of Guam. Within each region four school systems were asked to volunteer to evaluate a total of 20 teachers who would be proportionally randomly selected from one elementary school (8 to 10), one middle school (4 or 5), and one high school (5 to 7).

Answers:

Basically, there are three major considerations we should use in evaluating a measurement tool: validity, reliability, and practicality. We can improve the generalization of the study by standardizing the conditions under which the measurement takes place. The following is the evaluation for this scenario.

Validity

In this situation, the researcher should consider the validity as follows:

Internal validity:
Particularly, instrument (this threat results from changes, between observations, in the measuring instrument or observers), reactivity (subjects act differently when they realize they are subjects in the study, which is called "Hawthorne effect"), and privately held hypotheses and demand (subjects believe in the way researchers want them to behave or to make the researcher happy).

External validity:
Population validity (even though the subjects are randomly selected, it may not be the representative to the population) and ecological validity (the findings cannot be generalized to all contents of school level, place, and task; tat is, all have different characteristics).

Statistical conclusion validity:
In this case, sample schools are only four and 20 teachers as subjects are small. The small sample size in this case may cause low power for the study.

Construct validity:
In attempting to determine construct validity the researcher associates a set of other propositions with the results received from using the measurement tool. If the measurements on the devised scale correlate in a predicted way with these other propositions, then the researcher can conclude that there is some construct validity.

Control and causation and the use of randomization

The major method of controlling subjects bias or the effect of demand characteristics is to limit the subjects' knowledge about the general purpose of the study and about the hypotheses and variables being investigated. If the researcher wish to compare the two approaches in order to evaluate teacher effectiveness, the best technique is to assign the subjects randomly. In short, randomization is the optimum way to ensure group equivalence. Randomization also tends to uphold the internal validity of the study because it tends to assure that the samples are roughly similar in terms of subjects characteristics.

3 PROBLEM

3-1 The following incomplete ANOVA summary table is based upon a study in which the researcher used 150 observations with 25 students being randomly assigned to each of the six treatment groups. For this analysis he used IQ as a covariate, in which the common regression slope was .60 (r_{xy}=. 75). Complete the appropriate ANOVA summary table and indicate which are significant. Would the results be different without the covariate?

Source	SS	df	MS	F
Covariate	17360			
A	2000	5		
S/A	11440			
Total	30800	149		

Table 1

Source	SS	df	MS	F
Covariate	17360	1	17360	217*
A	2000	5	400	5*
S/A	11440	143	80	
Total	30800	149	206.71	

$^*p < .05$
F (.05, 1, 143) = 3.84 < 217, then reject H_o
F (.05, 5, 143) = 2.21 < 5, then reject H_o

Table 2

Source	SS	df	MS	F
A	2000	5	400	2
S/A	28800	144	200	
Total	30800	149		

$p > .05$
F (.05, 5, 144) = 2.21 > 2, then do not reject H_o

As illustrated in the Table 1, the effect of treatment is statistically significant with IQ scores (covariate) at the alpha level of .05. On the other hand, the effect of treatment is statistically nonsignificant without IQ scores (see Table 2).

Covariance is an intermediate figure on the way to finding the correlation coefficient. IQ scores and methods of treatment have a fairly strong positive correlation (r = .75). Because the correlation is the ratio of the covariance to the product of the standard deviations of X and Y. The steeper the slope, the larger the change in Y for a given change in X. In this case, each additional unit of X is associated with .60 additional units of Y.

3-2 *Below is given an incomplete ANOVA summary table. First complete the table as if all factors were fixed (as SPSS would). Then re-write the table, as it should be for C being a random factor and A and* ***B*** *being fixed factors.*

All Factors Fixed					Factors A & B Fixed, C Random			
Source	SS	df	MS	F	SS	df	MS	F
A	60	1						
B	198	2						
C	300	5						
AB	52	2						
AC	75	5						
BC	110	10						
ABC	50	10						
Within	1440	144						
Total	2285	179						

All factors fixed

Source	SS	df	MS	F	
A	60	1	60	6.00*	α = .05
B	198	2	99	9.90*	F (.05, 1, 144) = 3.84
C	300	5	150	15.00*	F (.05, 2, 144) = 3.00
AB	52	2	26	2.60	F (.05, 5, 144) = 2.21
AC	75	5	15	1.50	F (.05, 10, 144) = 1.83
BC	110	10	11	1.10	
ABC	50	10	5	.50	Thus, all main effects are
Within	1440	144	10		statistically significant at
					the alpha level of .05.
Total	2285	179	12.77		

* $p < .05$

Factors A & B fixed, C random

	SS	df	MS	F		
A	60	1	60	60/15	4.00*	
B	198	2	99	99/11	9.00*	Thus, in addition to all
C	300	5	150	150/10	15.00*	main effects, A x B
AB	52	2	26	26/5	5.20*	interaction is also statistically
AC	75	5	15	15/10	1.50	significant.
BC	110	10	11	11/10	1.10	
ABC	50	10	5	5/10	.50	
Within	1440	144	10			
Total	2285	179	12.77			

* $p < .05$

TM032433

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

Reproduction Release

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Research Design and Analysis: Examples of Questions and Answers

Author(s): Yukiko Inoue

Corporate Source: University of Guam

Publication Date: 2001

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced i ERIC system, Resources in Education (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, an Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following noti

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign in tl

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker s
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY SAMPLE TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY SAMPLE TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)	PERMISSI DISSEMIN MICROFICHE O SAMPLE TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
Level 1	Level 2A	Level 2B
X		
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g. electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche, or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Signature: Yukiko Inoue	Printed Name/Position/Title: Yukiko Inoue,PhD/Assistant Professor
Organization/Address: College of Education University of Guam UOG Station, Mangilao, Guam 96923	Telephone: 1-671-735-2447 — Fax: 1-671-734-4914
	E-mail Address: yinoue@uog.edu — Date: 2/6/2001

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
4483-A Forbes Boulevard
Lanham, Maryland 20706
Telephone: 301-552-4200
Toll Free: 800-799-3742
e-mail: ericfac@inet.ed.gov
WWW: http://ericfac.piccard.csc.com

EFF-088 (Rev. 9/97)