Chapter 16

Natural Environments: Observational Research

The last major class of research design that we'll examine is observational research. Designs for observational research may be used in laboratories, but they are most frequently used in natural settings. The major difference between observational research and the experimental research described in the previous two chapters lies in lack of manipulation of the independent or causal variable. The researcher relies on "natural processes" to provide different levels (variance) in the independent variable, and then simply observes simultaneous changes in both the independent and dependent variable.

The researcher may use an observational design because he has no other choice. Historical or retrospective studies offer no possibility for manipulating the independent variable, as the variables being observed have all had their values fixed sometime in the past. For example, if a researcher wishes to relate the childhood family communication patterns of adults to their current levels of marital adjustment, an experimental design (which would systematically manipulate childhood communication) is out of the question. Ethical issues may also prohibit manipulation of the independent variable. One could not deliberately create brain lesions in subjects in order to study the effects of left-hemisphere damage on nonverbal communication.

Because the system of variables being studied is allowed to operate with no intervention, the external validity of observational designs is usually very high. On the other side of the coin, the lack of control over the independent variable (because there is no direct, systematic manipulation of the causal variable by the researcher) and the uncontrolled effects of the environment on the dependent

variable (because these designs are frequently applied in the field), means that the internal validity of observational research is often very low. Lack of control for confounding variables (as we saw in Chapter 3) may mean that it is difficult to distinguish between a true relationship and a spurious or indirect one.

This difficulty is less critical in applied research situations where the primary objective is prediction of the outcome, rather than explanation of the theoretical process. For example, suppose a researcher's primary objective is to predict the readership of a new magazine. To do this, she examines the demographic characteristics of the likely audience, using variables like age, sex, geographic region, etc., and relates these to the readership of an existing magazine with similar content. She concludes that readers with demographic characteristics similar to the readers of the existing magazine will be likely to read the new magazine when it appears. By measuring the number of persons with these characteristics in the population to be reached by the new magazine, she can predict its initial readership.

The demographic variables describe the kinds of persons who read the existing magazine, and predict the number who are likely to choose to read the new magazine. But they will not explain why readers of a particular age, sex, or region choose to read either of the magazines. Perhaps older readers like the magazines because they have more experience with the topics addressed in the articles. If this is so, the relationship is spurious, even though there is covariance between age and readership. The true causal variable is amount of experience that the reader has had with the topics covered in the articles. If all readers were the same age, those with more experience with the topics would be more likely to read the magazine. But if the objective is to predict the readership in the real world, and since the real world contains readers of differing ages, this finding may not be of particular interest to the applied researcher.

But the scientist who wishes to go beyond simple prediction, and explain phenomena with observational designs must take pains to eliminate as many of the explanations that compete with the hypothesized relationship. The confidence which can be placed in the conclusions of an observational study is directly proportional to the elimination of competing explanations or spurious relationships.

It is never possible to eliminate all competing explanations for the observed changes in the dependent variable, so there is always the possibility that the covariance between two observed variables is spurious. Some people take this fact as evidence for the superiority of the experimental design, where spurious relationships can be eliminated by either manipulated control or randomization. But experiments frequently suffer from poor external validity, so there is a definite trade-off that the researcher must consider. A well-designed observational study which measures relevant confounding variables and uses statistical control to eliminate their effects could well be superior to a tightly controlled experiment which does not generalize to the real world.

Typical Observational Research Designs

The basic research designs are very simple compared to the experimental designs outlined in the previous chapters. Fundamentally, the researcher simply observes the naturally occurring values of the independent and dependent variables, and uses statistical methods to find out whether they covary. If the observation can be done without intruding on the ordinary behavior of the research subjects, the false responses that might be introduced by the manipulative research procedures are removed. This improves the external validity.

But the problems with the internal validity of observational research that we mentioned above mean that the research designs are very susceptible to confusion between the true effect of an independent variable and the spurious effects of other uncontrolled variables. Table 16-1 shows some threats to validity for the observational designs. If you compare it with the other tables for research designs in the previous chapters, you'll see that there are more internal validity problems and fewer external validity problems.

The research designs described here are both one-time-point, or "snapshot" designs. This is similar to the post-only experimental design described in the previous chapters, but differs from the two-time-point pre-post designs. Single time-point measurement is not susceptible to measurement learning effects, sensitization, history and maturation, subject mortality, and other threats to validity that occur over time. Multiple time point observational research studies are somewhat more

Table 16-1   Threats to Validity: Observational Research Designs

Simple Observational	*R:*	X	Y	(Observed Group)
		A		
		B		
		C…		[Control variables]
Group Comparison Observational	*S:*	X	Y	(Observed Group)
		A		
		B		
		C…		[Control variables]
	S:	X	Y	(Control or Comparison Group)
		A		
		B		
		C…		[Control variables]

	Simple Observational (a)	Group Comparison Observational (b)
INTERNAL VALIDITY		
Instrumentation Reliability and Validity	?	?
Selection (Sampling)	–	+
Instrument Obtrusiveness	?	?
Manipulation Effectiveness	–	–
History	?	?
Maturation	?	?
Measurement Sensitization	–	–
Measurement Instrument Learning	?	?
Measurement Instrument Instability	–	+
Subject Mortality	–	–
Subject Fatigue	?	?
Treatment or Measurement Order Effects	?	?
EXTERNAL VALIDITY		
Representative Sample	–	–
Reactive Effects of Setting	–	–
Multiple Treatment Interference	–	–
Reactive Sensitization	–	–
Subject Mortality	–	+

complex. These designs are covered in the next chapter. However, some of the threats to validity, like history and maturation, may confound onetime-point results, because the outside variables involved in history or maturation may produce spurious effects on the dependent variable.

The two major single-time-point designs are the simple observational design and the group comparison observational design. They differ in one fundamental way. In the simple observational design, a random representative sample is drawn from the population, and both independent and dependent variable are measured. Only a relationship hypothesis can be tested in this design. The statistic used to test the hypothesis will quantify any systematic covariance between observed levels of the independent and dependent variables. This might be a contingency table for nominal variables, or a correlation coefficient for interval or ratio variables (see Chapter 19).

In the group comparison observational design, the sample is divided into two or more groups according to the measured value of the independent variable. A comparative hypothesis is tested in this design, using a statistic that contrasts the groups on the observed levels of the dependent variable. Z- or t-tests, or Analysis of Variance F-ratios might be used to test these hypotheses (again, see Chapter 19 for an outline of these statistical tests).

There is a subtle, but critical, difference between the effects of random assignment of subjects in an experiment and the random selection of observations in an observational study. In an experiment, randomization insures that the effects of outside factors are approximately equivalent among all the experimental groups, and thus eliminates any systematic effects of these outside factors. But in an observational design, the random assignment is used to insure that the influence of the factors is the same in the study as it is in the real world. Randomization in observational designs will not remove the influence of external factors, but it will insure that the external validity of the study is good.

The group comparison design also uses random selection of observations, but the assignment to comparison groups is then done on a non-random basis. This non-random assignment means that the comparison groups may differ on factors other than the variable used for the assignment. For example, suppose the independent variable X is Age, and the dependent variable Y is the amount of usage of electronic mail. If the randomly sampled observations are assigned to two groups on the basis of Age (Young and Old), the groups will probably be unequal on other variables that are associated with age. The older group will probably have a higher income (an outside variable A). This could affect electronic mail usage, since an individual with a higher income could afford the cost of the service, is more likely to be a manager or technical worker to whom quick communication is necessary, etc. In order to avoid confounding the effects of income with those of age, the income levels in both comparison groups must be measured and statistically controlled. Statistical control is as important in comparative designs as it is in simple observational designs.

Observational measurement is often done "unobtrusively." This means that the units of analysis (individuals, groups, organizations, etc.) are not affected by the process of observation. In fact, they may not be even be aware that the research is being conducted. When this is the case, the effects of the research setting, like reactions to the laboratory setting, social demand responses, and especially measurement instrument obtrusiveness, are removed. Some typical procedures for unobtrusive measurement are covered in Chapter 18.

In either observational design, it is crucial that all confounding variables be measured and controlled statistically. Statistical control is the primary form of control exerted in observational designs.

Statistical Control Requirements

As we mentioned in Chapters 4 and 13, statistical control can be used as a partial solution to the problem posed by spurious relationships. But statistical control carries a large requirement: recognition, definition, and measurement of all relevant variables which might confound the relationship between the independent and dependent variables. Since we can never be sure that we've recognized all the confounding variables, we can never be absolutely sure that the covariance we find between the independent and dependent variables is not spurious.

Statistical control also requires more complicated statistical analysis procedures than we've discussed so far. Instead of dealing with just two variables (the independent and dependent variables), the researcher might have to account for the effects of dozens of confounding independent variables which must be controlled. In essence, the confounding variables must be introduced into

the research design as additional independent variables.

In plain words, multivariate statistical procedures are required. This opens a wide and sometimes scary world where data wizards and gurus lurk and perform feats of dazzling inference (and sometimes statistical sleight-of-hand). But all the wonderful technology and arcane vocabulary exist only to serve one purpose: control of confounding effects, so that the hypotheses that the researcher creates from theoretical reasoning can be tested and generalized.

Statistical control can sometimes be done simply by partitioning the data into subsamples in which the confounding outside variables are held constant. In our study of use of electronic mail, we could control for the effect of income by separating the sample into several income groups, like Low, Middle, High, and examining the relationship between Age and E-mail use within each constant income group. This is simple, but requires a large number of observations, particularly when the number of outside variables that have to be held constant is large. More complex statistics like partial correlation and regression that do the job of controlling the effects of outside variables are usually more efficient. But the also require measurement of the concepts at the interval or ratio level.

Multivariate statistical control of nominal and ordinal variables is possible, but it is involved. We'll mention some of the multivariate procedures that can do the job of statistical control at different levels of measurement in Chapter 19.

The prescription for research design of observational studies is clear. Define, at the highest level possible, all the relevant variables that your theoretical reasoning says might affect the dependent variable. Then measure them at this high level (ratio or interval, if possible, rather than nominal or ordinal). The complexity of the required multivariate statistical procedures will be reduced by higher levels of measurement.

Examples of Observational Research in Communication

The following examples illustrate observational designs applied to different units of analysis. The first example is a study conducted with the individual as the unit of analysis.

Example of Observational Research Using the Individual as the Unit of Analysis: Gerbner's "Mean World" and Cultivation Analysis

The Gerbner study described in Exhibit 16-1 used a simple observational design. The variables in the study are defined as occurring within the individual: the independent variable is the amount of time spent viewing television, and the dependent variable is the similarity between the person's view of social reality and the view shown by television. The independent variable is not manipulated. It is observed directly by asking the child viewers how much time they spend with television.

Many factors may work to produce differences in the amount of time spent viewing. This makes random selection of the observations very important, so that sample bias is not introduced, because the factors that affect viewing time may also affect the viewer's perception of social reality. This is why Gerbner's group measured and statistically controlled for the viewer's sex, age, socioeconomic status, IQ, etc. Each of these could be a confounding variable, and the effect of each must be removed before the researcher can attribute differing social views to the time spent with television.

Since the relationship between viewing time and having a social view which corresponded to that provided by television persisted when the competing explanations for this relationship were controlled, Gerbner concluded that television was cultivating a view of the world as being a much more violent place than it was in reality. Gerbner's critics, however, pointed to other unmeasured variables as possible confounding sources of the relationship. Because they were unmeasured, they could not be statistically controlled. This weakens somewhat the confidence which can be placed in the findings.

On the positive side, the viewing studied by Gerbner took place under natural conditions. Exposure to programming occurred before the child knew that a research project was being con-

EXHIBIT 16-1 Observational Research: Gerbner's "Mean World" and Cultivation Analysis

General Topic

Gerbner and associates developed a theory of media effects that states that media "cultivate" a particular view of society which may differ from reality. By providing a distorted picture of society, the media cultivate a similarly distorted sense of social reality in the audience members.

In a series of studies, Gerbner used content analysis of television programming to describe the kinds of situations and social behaviors that characterized the "television world." Content analysis is a technique that uses a message unit of analysis, like the "program", rather than an individual unit of analysis. An example of a content analysis is described later in this chapter.

Gerbner's research team related media content to perceptions of social reality, in order to test the idea that the media affect the audience's general view of their society.

Hypotheses

The cultivation hypothesis for television can be stated in general terms like this:

H1: The greater the amount of time spent viewing television, the more similar the person's view of reality will be to the view presented by television programming as a whole.

This hypothesis can only be tested when the events shown on television differ in frequency from those which occur in reality. If there is no difference, the viewer's perception of reality could be formed either by direct experience or by television viewing.

Operational Definition and Methods

Gerbner found a number of differences between the media depictions and the real world frequency of occurrence of acts such as violent crimes. Using this information, his research team constructed a number of questions that had a "TV an-swer" and a "Real World answer" and asked a sample of children to respond. A typical question might look like this:

During any given week, what are your chances of being involved in some kind of violence? About 1 in 10 or about 1 in 100.

The first answer was the "TV answer" as acts of violence appeared in programming far more frequently than they appear in reality (the actual probability of being involved in violence was about 1 in 200). By asking a number of these questions, and adding up the number of times the viewer gave "TV Answers," a variable which measured the overlap between the viewer's view of reality and the (false) TV depiction of that reality was constructed.

Time spend viewing television was measured directly by asking viewers to estimate their amount of viewing.

Results

Table E16-1 summarizes one of the tests of the general hypothesis. In Table E16-1, a partial correlation coefficient is used to describe the size of the relationship between viewing time and the number of TV answers given. (See Chapters 11 and 12 for details about correlation coefficients). A partial correlation coefficient is a version of the Pearson correlation coefficient which describes the size of a relationship in which the effects of other confounding variables are held constant. This allows the researcher to eliminate these variables as potential causes.

Like the correlation coefficient, partial correlation coefficients range from -1.0 (a perfect negative relationship) to +1.0 (a perfect positive relationship), with 0.0 indicating a null relationship. Table E16-1 shows that there is a significant relationship between the amount of viewing and the tendency of children to give "TV answers," rather than real world answers. The relationship persists even when the age, IQ, and sex of the child, the amount of newspaper reading done by the child, and the family's socio-economic status are

EXHIBIT 16-1 cont.

held constant. In other words, all differences due to these variables are removed from the estimate of the size of the relationship, and it still is positive enough that we can conclude that it is very unlikely that this result occurred simply by random sampling error. It appears that TV viewing is related to a child's incorrect view of reality that is consistent with the incorrect view presented by television.

Gerbner, G., Gross, L., Eleey, M.F., Jackson-Beeck, M., Jeffries-Fox, S., and Signorielli, N. (1977). TV violence profile #8: The highlights. *Journal of Communication*. 27, 176-207.

Gerbner, G., Gross, L., Eleey, M.F., Jackson-Beeck, M., Jeffries-Fox, S., and Signorielli, N. (1978). Cultural indicators: Violence profile no. 9. *Journal of Communication*. 28, 176-207.

Table E16-1   Relationship Between Amount of TV Viewing and Number of "TV Answers" to Questions About Violence and Law Enforcement Given by Children

Variable Controlled	Partial Correlation	Significance
None	.16	$p < .001$
Sex of Child	.15	$p < .001$
Age of Child	.16	$p < .001$
Newspaper Reading	.16	$p < .001$
Father's Education (Socio-Economic Status)	.15	$p < .001$
IQ	.13	$p < .001$
ALL CONTROLS	.12	$p < .004$

N = 466 children in New Jersey

ducted, so the research and measurement procedures could not have modified the viewing conduct of the research subject. Viewing was done in the home, not in a laboratory, so all the real world distractions, conversations, parental restriction on program choice, etc. were present. This gives credence to the claim that the relationship found in the research study is also present in the general population. A laboratory experiment would have controlled both the amount of viewing, and the conditions under which viewing took place. This control would have strengthened the conclusions, as many additional alternative explanations could have been ruled out, but it would have also been subject to the criticism that the effects observed would only occur under artificial laboratory conditions and are not present in the real world.

The Gerbner study also illustrates an common and important situation in communication research. The television programming is described with a content analysis that uses a message unit of analysis. As an example, the results of the content analysis found that 64% of TV characters were involved with violence. How can this finding about messages be related to the individual viewers, who are the primary units of analysis?

EXHIBIT 16-2   Observational Research: Social System Units of Analysis: Changes in Group Communication Relationships Over Time

General Topic

Coorientation variables describe some aspects of the relationship between two or more persons. The common variables are accuracy, agreement, and congruency (defined below). In the simplest form, they are obtained by comparing the answers to similar questions asked of two persons (we'll call them A and B) who have communicated about some topic (called X).

The amount of communication among the group members is expected to affect the coorientation variables. This study investigated the effects of communication over a time period in a small group of students working on laboratory assignments. It was essentially methodological and exploratory, but some theoretical propositions were tested.

Units of Analysis

Since the variables are defined by the relationship between the two sets of answers, the unit of analysis is not the individual, but is the dyad, or pair of participants in the communication. A and B can also be defined as social systems rather than individuals. In this research A was an individual, but B was defined as the "other members of the group," which is a social system unit. Treating a social system as if it is a composite individual is called "reifying" the group.

Definition of Variables

Accuracy is defined as the degree of similarity between the way that participant A describes participant B's view of some topic X, and the actual way that participant B views X. If A can describe B's view, A is accurate.

Agreement is the similarity between A's view of X and B's view of X. If both participants see X similarly, they agree.

Congruency is similarity between participant A's view of X, and participant A's description of participant B's view of X. If A *thinks* B views X the same way that A views it, there is congruency (even if A is inaccurate, and B actually views X differently).

Hypotheses

There were several issues addressed in this study. We will focus on one: the effect of the amount of communication on the three coorientation variables.

Increases in the amount of communication is expected to improve accuracy, as both members of the communication are more able to understand the other's perceptions. More communication should also increase the amount of agreement, as group members influence each other toward some common view of X. If the preceding two conditions are correct, communication would be expected to improve congruency, too. More communication will lead to more agreement, and to a more accurate perception of this agreement. These changes, in turn, increase congruency.

In this study, time is used as a surrogate variable for amount of communication (a surrogate variable is one that is correlated very strongly with an unmeasured variable, and is used in place of the unmeasured variable, usually because it is more convenient to measure). As time increases, it is presumed, so does the amount of communication among group members. Stated more formally in terms of time, the hypotheses are:

H_1: As time increases, so does the amount of accuracy in the dyad.

H_2: As time increases, so does the amount of agreement in the dyad.

H_3: If H1 and H2 are correct, as time increases, so does the amount of congruency in the dyad.

EXHIBIT 16-2 cont.

To measure congruency, person A's answer to the first question (5) is subtracted from persons A's answer to the second question (6), and the difference subtracted from 7:

CongruencyA = 7 - (6 - 5) = 6

A 12-member group which met weekly for 15 weeks was used to provide data for testing the hypotheses. At each group meeting, each member was given 16 to 17 items similar to the ones shown above. The items were varied from week to week, to minimize sensitization of the group members to particular attributes, but all attributes were measured several times during the 15 week period. Average values of accuracy, agreement, and congruency at 10 time points over the 15 week period were computed from the scales for a number of different attributes. An analysis of variance (see Chapter 19) was used to test the relationship between time (the session number, 1 to 10) and each of the coorientation variables.

Results

As hypothesized, accuracy, agreement, and congruency about the group attributes increased with time. Table E16-2 shows these results. The F-value, with its associated degrees of freedom, is the inferential statistic used to test for the significance of the relationship. It is related to the z-statistic described in Chapter 12, and, like that statistic, it is used to describe the probability that the relationship is due to sampling error. Table E16-2 shows that the probability of the relationship between time and changes in the coorientation variables being due to chance (random sampling error) is less than 5%.

Since the total amount of communication within the group also increased with time, it was concluded that communication produces improvements in the accuracy, agreement, and congruency of group members.

Steeves, H. L. (1984). Developing coorientation measures for small groups. *Communication Monographs.* 51, 185-192.

Table E16-2   Relationship of Time to Change in Coorientation Variables

Coorientation Variable	F-value	Degrees of Freedom	Significance
Accuracy	9.04	1;1220	$p < .05$
Agreement	5.17	1;1297	$p < .05$
Congruency	5.40	1;1219	$p < .05$

The common way is by introducing the concept of individual exposure to the message. In the Gerbner study this was done indirectly, by measuring the amount of time spent watching all TV programming. It was inferred that more time spent watching ("being exposed to..") all programming would automatically lead to more exposure to depictions of violence. In other studies of media violence, this exposure measurement has been done more directly. A researcher might count the acts of violence in a number of programs (the message unit of analysis), then ask the individual which programs he/she regularly views. A "violence viewing" score for the individual could then be constructed by adding the average number of acts of violence in each of the programs to which the subject reports being exposed. The violence viewing score is at the individual level of analysis, so it can be related to other individual level variables, such as attitudes or aggressive behaviors.

It is not possible to relate directly a variable with one unit of analysis to a variable with another unit of analysis. One must always use some reconceptualization device like "exposure" to bridge the different levels.

EXHIBIT 16-3   Observational Research: Message Units of Analysis: News-
paper Coverage of Political Campaigns

General Topic

The basic research question con-
cerned the fairness of the coverage which
the leading newspapers in the U.S. gave
to the 1968 presidential campaign. This
campaign was conducted at a time when
the electorate was polarized over the Viet-
nam War issue and other social issues
which arose from the Civil Rights move-
ment and the Johnson administration's
War on Poverty. George Wallace's Ameri-
can Independent Party challenged the
Republican candidate, Richard Nixon, and
the Democratic candidate, Hubert
Humphrey.

Stempel had conducted content
analyses of newspapers during the 1960
and 1964 campaigns. He found very little
difference in the amount of coverage given
to each of the two major party candidates
in that election, even though the pattern
of editorial endorsements for president
was disproportionally for the Republican
candidate in 1960 (Richard Nixon) and for
the Democratic candidate in 1964 (Lyndon
Johnson).

Stempel did not analyze the "tone"
of the coverage. Rather, he measured the
amount of space devoted to stories about
each candidate. The addition of the
Wallace candidacy in 1968, coupled with
some very passionate feelings about the
state of the country and some very impor-
tant differences among the candidates,
raised real questions about the nature of
media coverage.

Hypotheses

No formal hypotheses were stated.
The research question about fairness in
amount of coverage did imply a null hy-
pothesis, however: There is no difference
among the mean amount of coverage
given to the Democratic and Republican
candidates. The third-party candidate was
not expected to have an equivalent
amount of coverage. An additional re-
search question concerned the relation-
ship between an editorial endorsement by
a newspaper, and the amount of coverage

given that candidate. The implied research
hypothesis is this: The amount of news-
paper coverage will be higher for en-
dorsed candidates than for unendorsed
candidates. Finally, a research question
about the change in coverage over time
produced this implied nondirectional null
hypothesis: There will be no difference in
amount of coverage in this election, as
compared to past elections.

General Procedure

The newspaper (a aggregate mes-
sage unit) was the unit of analysis.
Amount of space in column inches (one
inch of type in a standard newspaper col-
umn width) was measured for each story
in the newspaper. The editorial endorse-
ment made by each newspaper was also
recorded. All stories from September 3 to
November 4, 1968, in 15 major newspa-
pers were measured.

Results

The results were presented as sum-
mary tables of descriptive statistics. A
typical summary is shown in Table E16-3.
The Democratic and Republican candi-
dates received very similar amounts of
coverage, and the differences from an even
split were not significant. There was very
little change in the percentages from elec-
tion to election, so no evidence for changes
in coverage proportions due to a chang-
ing political climate or the entry of a third
party were seen. However, analyzing the
endorsements and the amount of cover-
age from all three elections showed that
newspapers were much more likely (29
out of 37 times) to give more coverage to
candidates whom they endorsed. The dif-
ference in coverage between the endorsed
and unendorsed candidate in these cases
was small, but the pattern of giving more
space to the candidate of choice was sta-
tistically significant ($p < .001$).

The general conclusions were that
the coverage between the major parties
was well balanced. However, the political

EXHIBIT 16-3 cont.

Tpreference of the newspaper apparently was associated with a tendency to cover the endorsed candidate slightly more than the unendorsed candidate. This was considered a covariance relationship, and no causal inference was drawn.

Stempel, G. (1969). The Prestige Press Meets the Third-Party Challenge. *Journalism Quarterly*. 46(4), 699-706

Table E16-3   Percentage of Space for Each Party, 1960, 1964, 1968

	1960		1964		1968		
	Dem	*Rep*	*Dem*	*Rep*	*Dem*	*Rep*	*AIP*
Atlanta Constitution	60.2	39.8	53.0	47.0	36.4	34.6	29.0
Baltimore Sun	52.8	47.2	52.2	47.8	41.5	39.5	19.0
Chicago Daily News	48.8	51.6	53.9	46.1	39.0	33.9	27.1
Chicago Tribune	40.8	59.2	34.4	65.6	37.0	45.9	17.0
Christian Science Monitor	41.5	58.5	48.9	51.1	37.2	32.9	29.9
Des Moines Register	45.8	54.2	50.5	49.5	36.9	42.1	20.9
Kansas City Star	52.2	47.8	54.4	45.6	36.0	37.1	26.8
Los Angeles Times	45.2	54.8	49.9	50.1	48.1	32.7	19.1
Louisville Courier-Journal	54.5	45.5	57.8	42.2	39.4	30.2	30.4
Miami Herald	51.3	48.7	59.4	40.6	39.3	35.0	25.7
Milwaukee Journal	54.4	45.6	51.0	49.0	42.7	37.5	19.8
New York Times	50.3	49.7	57.0	43.0	43.9	38.0	18.0
St. Louis Post-Dispatch	54.6	45.4	52.0	48.0	44.2	36.6	19.2
Wall Street Journal	47.2	52.8	45.2	54.8	41.4	37.6	21.4
Washington Post	51.9	48.1	53.5	46.5	44.3	35.9	19.9
Average	50.2	49.8	51.9	48.1	41.4	36.8	21.8

Dem=Democratic
Rep=Republican
AIP=American Independent Party

Example of Observational Research Using the Social System as the Unit of Analysis: Steeves' study on Coorientation in Small Groups

Unlike the Gerbner research, the Steeves study described in Exhibit 16-2 uses variables that require combined measurement of two or more individuals. The makes the unit of analysis the social system. It is not possible to assign an agreement score to a single individual, as agreement is defined as the relationship between measurements made on two individuals.

Like all observational studies, there was no manipulation of the independent variable. The amount of communication was allowed to take on natural values without researcher intervention.

The basic design was the simple observational study. Although the measurements were taken over a number of weeks, time is actually being used to represent the amount of communication. If there was a direct measurement of the amount of communication made within different groups, this would be a snapshot study. However, the measurement procedure introduces some of the problems and threats to validity that occur with over-time research. These are discussed in detail in the next chapter.

Steeves apparently concluded that there were no serious confounding variables which might produce a spurious relationship, as she did not use statistical control for any third variables in her analysis. This decision might be open to some criticism. But it is difficult to think of variables other than communication which would affect accuracy, agreement, or congruency. In the absence of plausible competing explanations for change in the dependent variables, the researcher can conclude that all relevant variables have been measured, and the likelihood of the relationship being spurious is not great.

Example of Observational Research Using the Message as the Unit of Analysis: Stempel's Study of the Prestige Press

Stempel's content analysis is a very common kind of observational study. It uses a group comparison observational design which is unobtrusive, as the coverage was measured retrospectively. Unlike the prototype design shown in Table 16-1, the units of analysis were not chosen randomly. Stempel essentially defined the prestige press as consisting of the newspapers which were measured. He thus conducted a census of the units, so no sample bias was introduced.

The independent variable (actually, a nominal factor) was the political party covered. Within each group defined by the levels of this factor, the dependent variable of amount of coverage was observed. Since this was a census, no statistical tests were required in Table E16-3. But the research question concerning the relationship between endorsement of a party's candidate and the amount of coverage given the party was subjected to a test. In this case, Stempel is treating the selected newspapers as a sample, and is implicitly generalizing to other newspapers. If he was not doing so, he would need no statistical test, as the difference observed in this census would have no sampling error, i.e., would be absolutely accurate.

As pointed out in the Gerbner study, a content analysis such as this that uses the message unit of analysis is frequently used in conjunction with a study using the individual unit of analysis. This linkage is one of the distinguishing hallmarks of communication research. Both message and individual qualities are observed and relationships between the two are explored. While a psychological study might concentrate on the individual, and a literary or textual analysis might concentrate on the message, good communication research designs frequently look at both.

Summary

In experimental research, the independent variable is systematically manipulated by the researcher so that it takes on different values, and the dependent variable is observed to see if it covaries. In observational research, the independent variable is not manipulated, but is allowed to take on different values "naturally," without the intervention of the researcher. Both the independent and dependent variables are measured and tested for covariance.

Because the phenomenon under study is allowed to occur naturally, the researcher has much less control over the process. This makes it harder to eliminate outside variables which might be

setting up spurious relationships between the independent and dependent variables. The observational researcher must rely on statistical control as the major tool to prevent this.

Statistical control, however, requires that all outside variables which might confound the relationship under study must be measured and used in the control procedure. Since it is difficult to state with full confidence that all confounding variables have been recognized, observational research designs are susceptible to a number of internal validity problems.

But observational designs are generally stronger in external validity than experimental designs. The lack of control over the operation of the phenomenon that decreases internal validity makes generalization to the real world easier. Adding to this external validity, observational designs are frequently used in field, rather than laboratory, settings.

There are two basic observational designs. The simple observational design requires measurement of the independent and dependent variables in a representative random sample. It is used to test relationship hypotheses.

In the group comparison observational design, sampled units are assigned to different groups according to observed (rather than manipulated) levels of the independent variable. Since this assignment is not random, the groups will differ not only on the level of the independent variable, but on any outside variables which are systematically related to the independent variable. This is unlike experimental research, where random assignment of units to experimental groups allows the researcher to assume that the groups are equivalent on all variables except the independent variable. Statistical control of outside variables is required in the group comparison observational design, as well as the simple observational design.

Observational designs must be used in retrospective research, and any research setting where manipulation of the independent variable is impossible for practical or ethical reasons. An observational design may also be preferable where the importance of external validity outweighs the additional difficulty of measuring confounding variables, and controlling their effects statistically.

References and Additional Readings

Cochran, W.G. (1983). *Planning and analysis of observational studies*. New York: Wiley.

Kidder, L. H. (1981) *Selltiz, Wrightsman and Cook's research methods in social relations*. New York: Holt, Rinehart and Winston. (Chapter 16: "Survey Research Design")

Webb, E.J, Campbell, D.T., Schwartz, R.D., & Sechrest, L. (1972). *Unobtrusive measures: Nonreactive research in the social sciences*. Chicago: Rand McNally. (Chapter 5, "Simple Observation").

Wimmer, R.D. & Dominick, J.R. (1987). *Mass media research: An introduction (2nd Edition)*. Belmont, CA: Wadsworth. (Chapter 6, "Survey Research").