**Section 1.2 Observational Studies versus Designed Experiments**

**Objectives**

1. Distinguish between an Observational Study and a Designed Experiment
2. Explain the Various Types of Observational Studies

***Objective 1: Distinguish between an Observational Study and a Designed Experiment***

Objective 1, Page 1

*Answer the following as you watch the video.*

*1*

Why is the Danish study mentioned in the video an observational study and not a designed experiment? **That is, in an observational study,**

**the researcher simply observes the behavior of the individuals**

**in the study and records the values**

**of the explanatory and response variables.**

2 Why is the “rat” study mentioned in the video a designed experiment and not an observational study **Because the researchers manipulated**

**the value of one explanatory variable, RFR exposure,**

**and controlled other values of explanatory variables,**

**such as temperature, humidity, food,**

**and so on, at fixed values, and there was random assignment,**

**the study is called a designed experiment.**

3.What is the response variable in each study, and what is the explanatory variable? In both studies the goal

was to determine if radio frequencies from cellphones

increase the risk of contracting brain tumors.

**Whether the individual contracted a brain tumor or not**

**was the response variable.**

**For the human study, the explanatory variable**

**of interest was cellphone usage or not.**

**For the rat study, the explanatory variable**

**was radiofrequency radiation, which**

**had three possible levels--**

**no RFR, GSM, or CDMA.**

***In research, we wish to determine***

***how varying an explanatory variable affects the value***

***of a response variable.***

Objective 1, Page 2

*Answer the following after watching the video.*

In research, we wish to determine how varying an explanatory variable affects … ***variable affects the value***

***of a response variable.***

1. What does an observational study measure? Does an observational study attempt to influence the value of the response variable or explanatory variable? Neither

OBJECTIVE 1, PAGE 2 (CONTINUED)

Explain how you would determine if a study is a designed experiment. Whether the explanatory variable values were manipulated or not.

Objective 1, Page 4

*Watch the video and answer the following.*

Why is the influenza study mentioned in the video an observational study and not a designed experiment? Observational. researchers, did not force the seniors to either get a flu shot or not.

But rather it was the seniors who decided that, hey I'm

going to go get a flu shot, or I'm

going to choose not to get a flu shot.

In addition, we as researchers did not

impose any other conditions on the study.

List some changes that could be made to investigate the effectiveness of the flu shot with a designed experiment. We could separate the seniors into two groups and measure more lurking variables and confounding variables. Such as Health, genetics, diet etc…

List some lurking variables in the influenza study. some lurking variables

in the influenza study-- age, health status, mobility

of the senior.

What are some variables (besides getting a flu vaccine) that may play a role in whether one contracts pneumonia or influenza? Genetics, and above lurking variables.

Objective 1, Page 4 (continued)

Define confounding in a study. So again, confounding in a study occurs

when the effect of two or more explanatory variables

are not separated.

Therefore, any relation that may exist

between an explanatory variable and the response variable

may be due to some other variable or variables

not accounted for in the study

What is a lurking variable? So a lurking variable is an explanatory variable

that was not considered in the study,

but that effects the value of the response variable.

**Do observational studies allow a researcher to claim causality**? No because in observational studies

we are never allowed to make statements of causality.

You cannot say the changes in an explanatory variable cause some

change in a response variable in observational studies.

You can only say that there's an association.

It requires designed experiments to make

statements of causality.

So observational studies do not allow a researcher

Objective 1, Page 7

List some reasons why an observational study would be conducted if causation cannot be claimed. Why would we ever conduct an observational study if we cannot claim causation? Because it is often unethical to conduct a designed experiment, also it has lower cost, greater timeliness, and a broader range of patients

Objective 1, Page 8

Define: Confounding variable confounding in a study occurs

when the effect of two or more explanatory variables

are not separated.

Therefore, any relation that may exist

between an explanatory variable and the response variable

may be due to some other variable or variables

not accounted for in the study.

Objective 1, Page 9

What is the big difference between lurking variables and confounding variables?

EXAMPLE

In designed experiments, it is possible to have two explanatory variables in a study that are related to each other and related to the response variable. For example, suppose Professor Egner wanted to conduct an experiment in which she compared student success using online homework versus traditional textbook homework. To do the study, she taught her morning statistics class using the online homework and her afternoon class using traditional textbook homework. At the end of the semester, she compared the final exam scores for the online section to the textbook section. If the morning section had higher scores, could Professor Egner conclude that online homework is the cause of higher exam scores? Not necessarily. It is possible that the morning class had students who were more motivated. It is impossible to know whether the outcome was due to the online homework or to the time at which the class was taught. In this sense, we say that the time of day the class is taught is a *confounding variable*.

confounding in a study occurs

when the effect of two or more explanatory variables

are not separated.

Therefore, any relation that may exist

between an explanatory variable and the response variable

may be due to some other variable or variables

not accounted for in the study.

EXAMPLE

In designed experiments, it is possible to have two explanatory variables in a study that are related to each other and related to the response variable. For example, suppose Professor Egner wanted to conduct an experiment in which she compared student success using online homework versus traditional textbook homework. To do the study, she taught her morning statistics class using the online homework and her afternoon class using traditional textbook homework. At the end of the semester, she compared the final exam scores for the online section to the textbook section. If the morning section had higher scores, could Professor Egner conclude that online homework is the cause of higher exam scores? Not necessarily. It is possible that the morning class had students who were more motivated. It is impossible to know whether the outcome was due to the online homework or to the time at which the class was taught. In this sense, we say that the time of day the class is taught is a *confounding variable*.

**The Difference between Lurking Variables and Confounding Variables**

The big difference between lurking variables and confounding variables is that lurking variables are not considered in the study (for example, we did not consider lifestyle in the pneumonia study) whereas confounding variables are measured in the study (for example, we measured morning versus afternoon classes).

So lurking variables are related to both the explanatory and response variables, and this relation is what creates the apparent association between the explanatory variable and response variable in the study. For example, lifestyle (healthy or not) is associated with the likelihood of getting an influenza shot as well as the likelihood of contracting pneumonia or influenza.

A confounding variable is a variable in a study that does not necessarily have any association with the other explanatory variable but does have an effect on the response variable. Perhaps morning students are more motivated, and this is what led to the higher final exam scores, not the homework delivery system.

The bottom line is that both lurking variables and confounding variables can confound the results of a study, so a researcher should be mindful of their potential existence.

***Objective 2: Explain the Various Types of Observational Studies***

. cross-sectional study.

And this is an observational study that collects information

about individuals at a specific point in time

or over a very short period of time.

Do smokers have a higher resting heart rate than nonsmokers?

You just go out and ask people.

Hey, are you a smoker or are you not a smoker?

That's how you segment your population.

And then, you would measure their heart rate.

This would be a cross-sectional study,

because it's over a very short period of time.

18. A case control study.

These studies are retrospective.

Retrospective basically means that you look back in time.

Case control studies look back in time

or require the researcher or to look at existing record, which

is also looking back in time.

In case control studies, individuals

that have certain characteristics

are matched with those that do not.

That's why it's case control, because with this smoker thing

we might say, well, how many packs of cigarettes

have you smoked on average each week for the past 10 years.

And we would match those folks with people

with similar demographics, characteristics

that did not smoke.

And then, we might compare the rate of lung cancer

in the two groups.

A case control study.

These studies are retrospective.

Retrospective basically means that you look back in time.

Case control studies look back in time

or require the researcher or to look at existing record, which

is also looking back in time.

In case control studies, individuals

that have certain characteristics

are matched with those that do not.

That's why it's case control, because with this smoker thing

we might say, well, how many packs of cigarettes

have you smoked on average each week for the past 10 years.

And we would match those folks with people

with similar demographics, characteristics

that did not smoke.

And then, we might compare the rate of lung cancer

in the two groups.

19. The negative thing about case control studies

is there's a memory issue here.

People might not remember how many packs of cigarettes

they smoked on average for the past 10 years.

And therefore, you might get bogus information out

of the study.

Or records might not exist or might be somewhat flawed,

which impacts the study.

But case control studies, none the less,

are better than cross-sectional studies,

because you have a matching going on.

You have a control group to compare

against your sort of research group, nonsmokers

versus the smokers.

But the negative, again, is that, looking back in time,

you often get erroneous information.

21. A cohort study first identifies a group of individuals

to participate in the study.

This is called your cohort.

And then, the cohort is observed over a period of time.

Over this period of time, characteristics

about individuals are recorded.

So these are called prospective studies,

because they are done going forward.

So basically, what I would do is something along these lines.

I would identify 5,000 individuals

and literally follow them over time, periodically

recording information about the individuals.

22. The advantage here is that you don't

have to rely on folks looking back in time.

And you don't have to rely on existing records,

because you are collecting the information over time.

The famous cohort study that exists

is the Framingham Heart Study, which

has been going on for years.

And in fact, is now in new generations

and because of this Framingham Heart cohort study,

a lot of new ideas and breakthroughs in medicine

have occurred.

23. The negative about these studies is

that they are extremely time intensive and labor

intensive and, therefore, usually very expensive.

Objective 2, Page 2

**Example 1** ***What Type of Study?***

Determine whether each of the following studies depict an observational study or an experiment. If the researchers conducted an observational study, determine the type of the observational study.

1. Researchers wanted to assess the long-term psychological effects of children evacuated during World War II. They obtained a sample of 169 former evacuees and a control group of 43 people who were children during the war but were not evacuated. The subjects’ mental states were evaluated using questionnaires. It was determined that the psychological well being of the individuals was adversely affected by evacuation. (Source: Foster D, Davies S, and Steele H (2003) The evacuation of British children during World War II: a preliminary investigation into the long-term psychological effects. Aging & Mental Health (7)5.)
2. Xylitol has proven effective in preventing dental carries (cavities) when included in food or gum. A total of 75 Peruvian children were given milk with and without xylitol and were asked to evaluate the taste of each. Overall, the children preferred the milk flavored with xylitol. (Source: Castillo JL, et al (2005) Children’s acceptance of milk with xylitol or sorbitol for dental carries prevention. BMC Oral Health (5)6.)
3. A total of 974 homeless women in the Los Angeles area were surveyed to determine their level of satisfaction with the healthcare provided by shelter clinics versus the healthcare provided by government clinics. (Source: Swanson KA, Andersen R, Gelberg L (2003) Patient satisfaction for homeless women. Journal of Women’s Health (12)7.)
4. The Cancer Prevention Study II (CPS-II) is funded and conducted by the American Cancer Society. Its goal is to examine the relationship among environmental and lifestyle factors on cancer cases by tracking approximately 1.2 million men and women. Study participants completed an initial study questionnaire in 1982 providing information on a range of lifestyle factors such as diet, alcohol and tobacco use, occupation, medical history, and family cancer history. These data have been examined extensively in relation to cancer mortality. Vital status of study participants is updated biennially. Cause of death has been documented for over 98% of all deaths that have occurred. Mortality follow-up of the CPS-II participants is complete through 2002 and is expected to continue for many years. (Source: American Cancer Society)

Type of Study?

A. Observational- Case Control

B. Experimental

C. Observational- Cross Section

D. Observational- Cohort

Objective 2, Page 3

It is not always possible to conduct an experiment. Explain why we could not conduct an experiment to investigate the perceived link between high tension wires and leukemia (on humans**). It is not ETHICAL to conduct some experiments.**

**me Concluding Remarks about Observational Studies Versus Designed Experiments**

Is a designed experiment superior to an observational study? Not necessarily.

* Because cross-sectional and case-control observational studies are relatively inexpensive, they allow researchers to explore possible associations prior to undertaking large cohort studies or designed experiments.
* It is not always possible to conduct an experiment. For example, we could not conduct an experiment to investigate the perceived link between high-tension wires and leukemia (on humans). Do you see why?

25. Have you heard this saying?

*"There is no point in reinventing the wheel."*

Here is how it applies to statistics: There is no sense expending energy to obtain data that already exist. If a researcher wants to conduct a study and an appropriate data set exists, it would be silly to collect the data from scratch.

For example, various federal agencies regularly collect data that are available to the public. Some of these agencies include the Centers for Disease Control and Prevention ([www.cdc.gov](http://www.cdc.gov/)), the Internal Revenue Service ([www.irs.gov](http://www.irs.gov/)), and the Department of Justice (<http://fjsrc.urban.org/index.cfm>). Another useful source of data is the General Social Survey (GSS), [www.gss.norc.org](http://www.gss.norc.org/), administered by the University of Chicago. This survey regularly asks "demographic and attitudinal questions" of individuals around the country.

26.   
A **census** is a list of individuals in a population along with certain characteristics of each individual.

27. The United States conducts a census every 10 years to learn the demographic makeup of the United States. Everyone whose primary residence is within the U.S. borders must fill out a questionnaire packet. The cost of obtaining the census in 2010 was approximately $5.4 billion; about 635,000 temporary workers were hired to assist in collecting the data.

Why is the U.S. Census so important? **The results of the census are used to determine the number of representatives in each state in the House of Representatives, boundaries of congressional districts, distribution of funds for government programs (such as Medicaid), and planning for the construction of roads and schools.** The first U.S. Census was conducted in 1790 under the direction of Thomas Jefferson. It is a constitutional mandate that a census be conducted every 10 years.

Is the United States successful in obtaining a census? Not entirely. Some individuals go uncounted due to illiteracy, language issues, and homelessness. Given the political stakes that are based on the census, politicians often consider how to count these individuals. Statisticians have offered solutions to the counting problem. If you wish, go to [www.census.gov](http://www.census.gov/); in the search box, type *count homeless*. You will find many articles on the U.S. Census Bureau's attempt to count the homeless. The bottom line is that even census data has flaws.

to claim causation, only association.