

Experiments

1000: Introduction

Suppose that you work for a major pharmaceutical company, and you have just discovered an over the counter herbal drink from Venezuela that appears to stop the growth of cancer. Suppose you have data in which there is an association between taking the drug and the remission of cancer. Should you stop all other research and try to extract the drug in the herbal remedy so you can synthesize and save thousands of lives and make millions of dollars?

Not quite yet. Why? Because, as you have now heard dozens of times, a simple association between two variables does not tell us whether one of the variables causes the other. In this situation, the association between taking the herbal remedy and cancer remission does not determine whether the drug actually causes the remission of the cancer, or whether there is some common cause of the two. For example, people who take the drug might be the 20% of cancer victims who are aggressive enough to seek out and try lots of different treatments, and perhaps its the spirit of determination that underlies their quest that boosts their own immune system and cures their cancer, not the herbal remedy. How do you, the researcher, go about determining whether the drug actually stops the growth of cancer?

In 1935, Sir Ronald Fisher, perhaps the most influential statistician in history, published a textbook called "Experimental Design" that addressed exactly this problem: how do we tell whether an association is due to a common cause, or to one variable's direct influence on the other. In this book, which is remarkably accessible even to laymen, Fisher describes what is still the most commonly used method for causal discovery in medical, social and behavioral science: **experiments with randomized assignment of treatment**.

This module introduces you to experiments generally and explains qualitatively why they work for causal discovery. The theory of experimental design can become quite complicated, but its core idea is quite simple. Here we try to make the core idea vivid with how causal graphs represent causation, how they produce association, and how interventions change causal graphs. We will then go on to discuss two important kinds of experiments: controlled experiments, and experiments with randomized assignment of the treatment. Since we will be using many concepts from the following modules, then if they are not fresh in your mind, take two or three minutes to review the Summary sections in:

- 1 Interventions
- 2 Causation to Association I: Causation to Unconditional Association

As important as they are, experiments are not always the best tool for causal discovery nor the only techniques available. As an example of why they aren't the perfect tool, experiments require that we can intervene on the causes under study. In many cases this is impossible. For ethical reasons we cannot do an experiment in which we infect humans with HIV virus to see if it causes AIDs. In many other cases, it is impossible for practical reasons: we cannot move the moon to see if it truly is the cause of the ocean's tides. This inability to perform experiments hasn't stopped our efforts at causal discovery in these cases, however, as we have still managed to discover that HIV causes AIDs and that the moon causes the tides with high confidence. There are also other techniques available for causal discovery. We cover non-experimental statistical techniques, such as statistical control and instrumental variables, in other modules. Here, we focus on the why and how of experimentation in general, and controlled and randomized experiments in particular.

[2000: The Power of Experiments](#)

[2100: The Problem of Alternatives](#)

As we discussed at length in the module on Causal Prediction vs Causal Discovery, and in the module on Confounding, associations underdetermine causation. Many causal graphs make the same predictions about associations. For example, all of the following causal graphs predict that **X** and **Y** are associated:

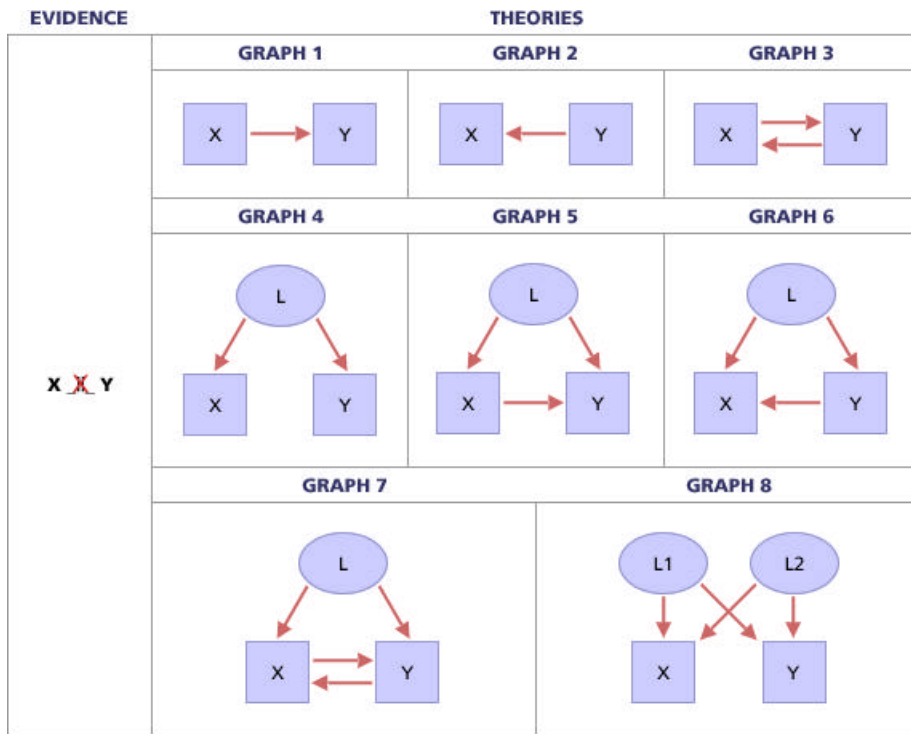


FIGURE 2100-1

And this is only a small fraction of all the graphs that make the same prediction. By adding some background knowledge, we can reduce this set. For example, if we know that **X** occurs before **Y**, then **Y** cannot cause **X**, and so we reduce the set to alternative 1, 4, 5, and 8, but we still don't have good evidence that **X** is a cause of **Y**.

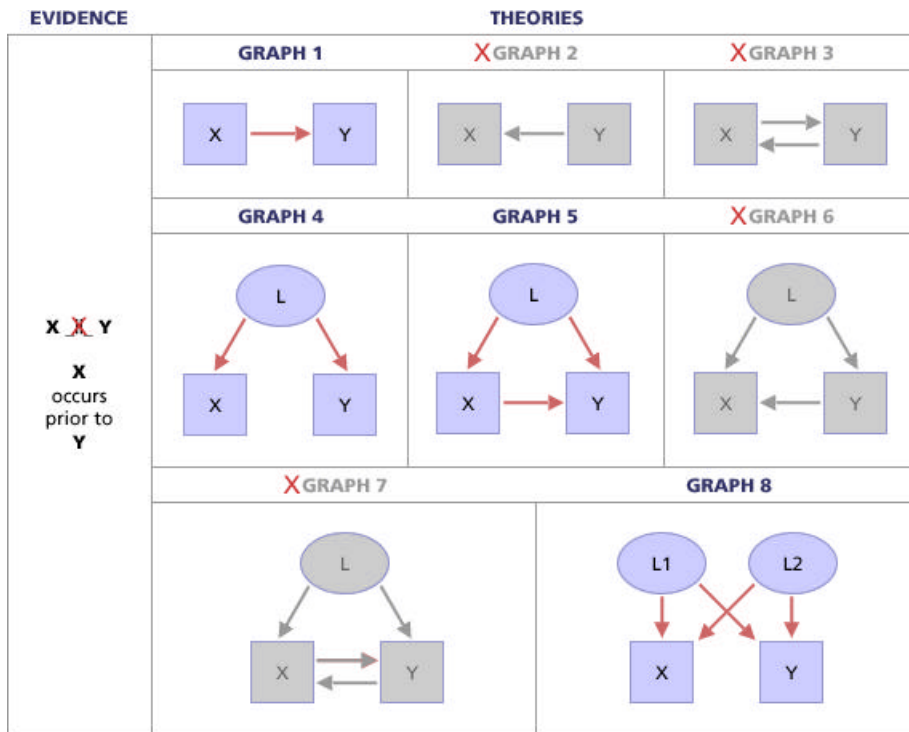


FIGURE 2100-2

< [A link to exercises in the interactive version of this module.](#) >

A Toy Example

The best answer to the problem of alternatives is experimental intervention. Consider the case of education and income. Suppose we check the U.S. Census, and extract data on education and income for all 35-year-old Americans. To be concrete, suppose for each 35-year-old we record their values on two variables: **EDUCATION** [High School or less, College or more] and current **INCOME** [Low: less than \$20,000, Medium: \$20,000-\$60,000, High: greater than \$60,000]. Since most people finish school (whatever the level) by age 30, we can assume that the **EDUCATION** variable almost always has its value by then. Since **INCOME** is measured when they are 35, we can assume that **INCOME** does not cause **EDUCATION** in this graph. (Note that **INCOME** might be a cause of **EDUCATION** for other people, such as young children; we're just assuming it isn't a cause for 35-year-olds.) Furthermore, suppose that in the Census data **EDUCATION** and **INCOME** are strongly associated. This association can only be explained by one of the following graphs:

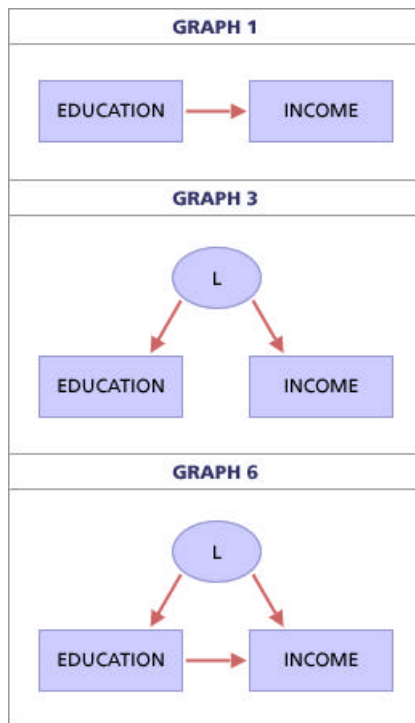


FIGURE 2100-3

where "L" stands for any unmeasured common cause of **EDUCATION** and **INCOME**, e.g., parental influence.

Now suppose we really want to know which of these theories is true, so we decide to perform an experiment in which we intervene to set the value of **EDUCATION**. We get a lot of money from a grant and perform a 25 year experiment that begins on 20,000 15-year-olds in which we don't let half of our subjects get any education after High School, and we make sure the other half at least finishes college. We then measure each person's **INCOME** when he or she reaches 35. Although unethical, fascist, and completely impractical, imagine that we could do it. What is the post-manipulation graph for each theory in the figure above? Remember that our intervention is ideal. Whereas parental influence might have influenced **EDUCATION** in the population before our intervention (the Census data), in our experiment it won't have any effect at all.

< [A link to exercises in the interactive version of this module.](#) >

An experiment in which we ideally intervene on **EDUCATION** leaves us with the following post-manipulation alternatives.

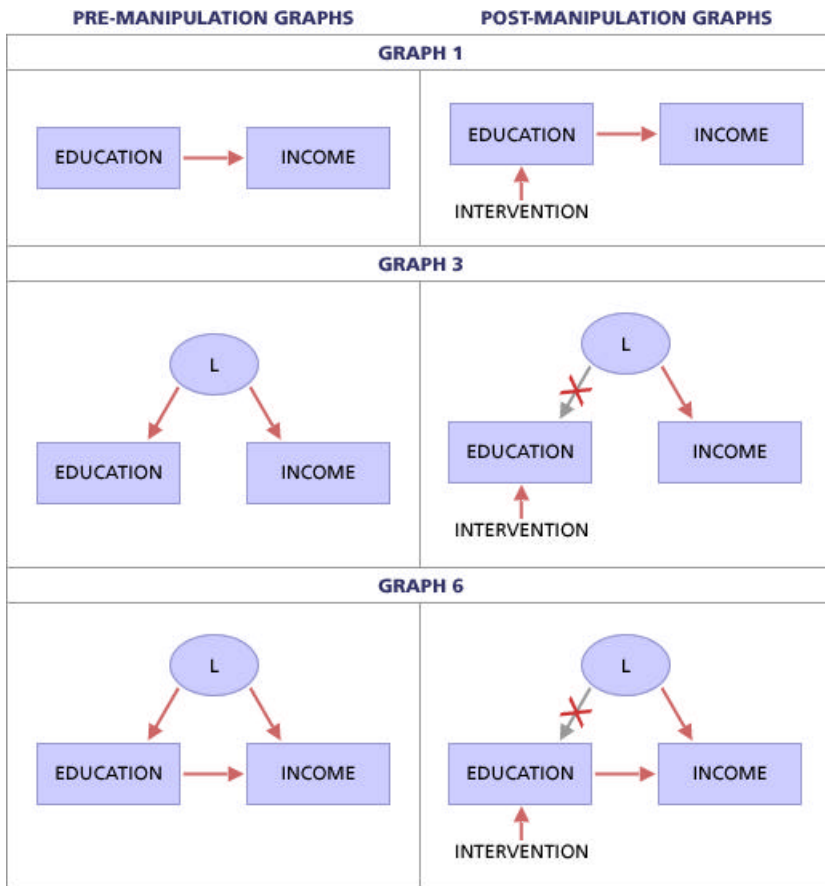


FIGURE 2100-4

[< A link to exercises in the interactive version of this module. >](#)

Although all three of our alternative theories predict that **EDUCATION** and **INCOME** are associated in the pre-manipulation setting (Census data), only alternatives 1 and 6 predict that **EDUCATION** and **INCOME** are associated in the post-manipulation setting (our experiment). **EDUCATION** is a cause of **INCOME** only in alternatives 1 and 6, so our intervention has allowed us to draw causal conclusions from associational data, even though we could not do so before the intervention.

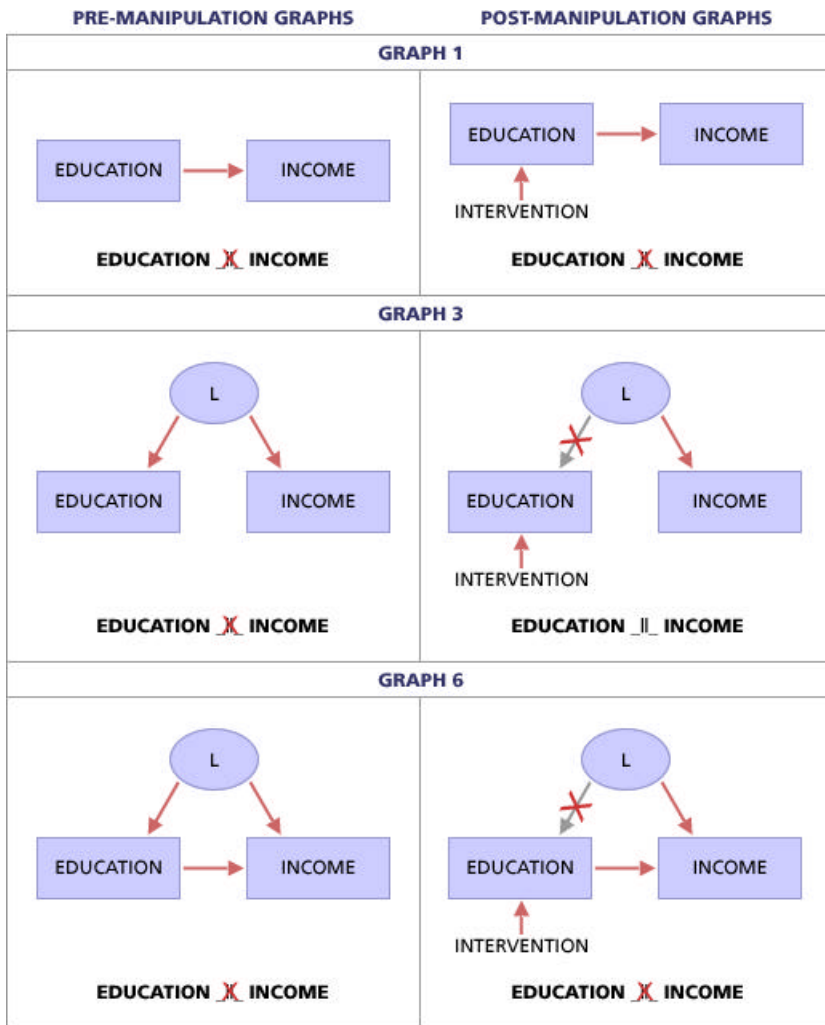


FIGURE 2100-5

< [A link to exercises in the interactive version of this module.](#) >

2200: The General Theory

Experiments in which we ideally intervene on a variable X eliminate all causal relationships in which X is an effect, and this change between the pre_ and post_manipulation setting affords us enormous scientific power in using association to discover causation. In this section we leave concrete examples behind to make the point generally.

Consider the situation among **X** and **Y** before conducting an experiment. Suppose we have no idea whether **X** occurs before **Y** or vice versa; we know only that **X** and **Y** are associated. Before conducting an experiment, the following alternatives might explain this association, where again "L" represents the existence of at least one unmeasured confounder:

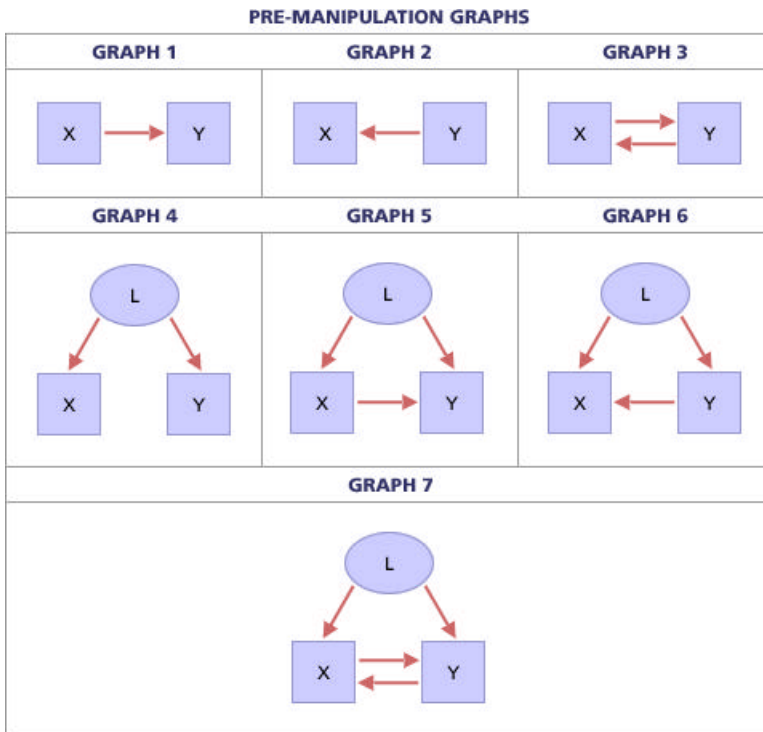


FIGURE 2200-1

Now suppose we conduct an experiment in which we ideally intervene on **X**. Here are the seven post-manipulated graphs.

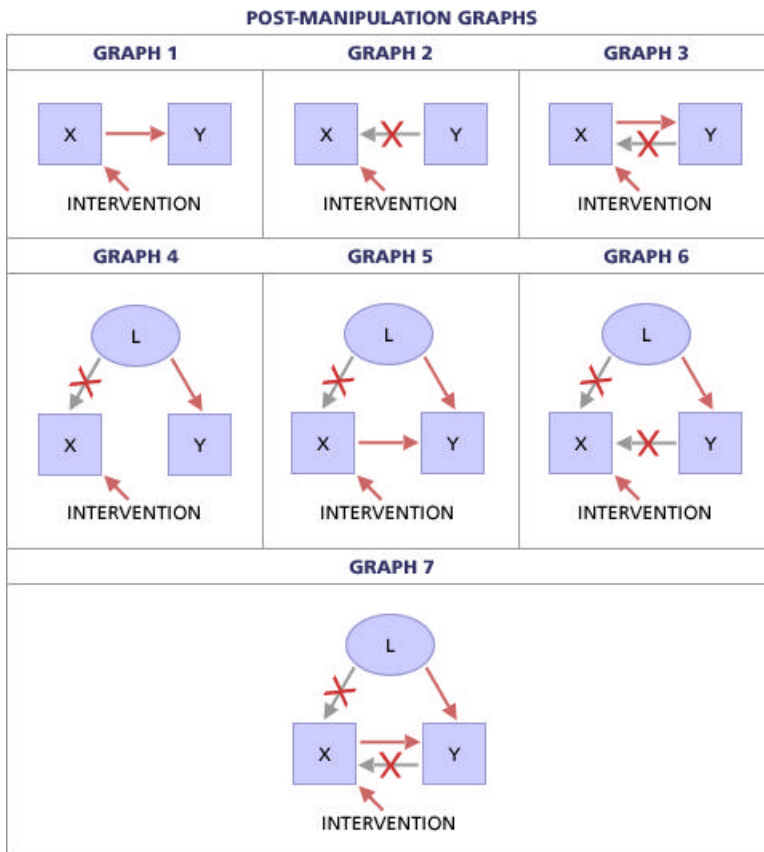


FIGURE 2200-2

< [A link to exercises in the interactive version of this module.](#) >

Experimental manipulation by ideal intervention on **X** is the best method of discovering whether **X** is a cause of **Y** because the ideal intervention eliminates all causal connections between **X** and **Y** except for one: **X** causes **Y**.

This is so important we'll repeat it in a different way. No matter how **X** and **Y** are causally connected in the natural, pre_manipulation graph, the only causal connection that can possibly hold between them after an ideal intervention on **X** is a causal path from **X** to **Y**.

Since **X** and **Y** can only be associated if they are causally connected, we can actually use the **data** from experiments to discover **causal** relationships. In other words, because of the how ideal interventions change a causal graph:

If **X** and **Y** are associated in an experiment in which we have ideally intervened on **X**, then **X** must be a cause of **Y**.

If **X** and **Y** are associated in passively observed data, then **X** and **Y** are causally connected, but we usually don't know how. This is why experiments are so powerful.

2300: Examples

2310: Finger Stains and Lung Cancer (again)

Recall the story of Burt and Russell's lung cancer prevention program in the module on Causation vs. Association. Having observed an association between **FINGER STAINS** and **LUNG CANCER**, Burt and Russell decided that an intervention which eliminated **FINGER STAINS** would also drastically reduce the incidence of **LUNG CANCER**. Unfortunately, here is the causal story behind smoking, finger stains, and lung cancer in the real world.

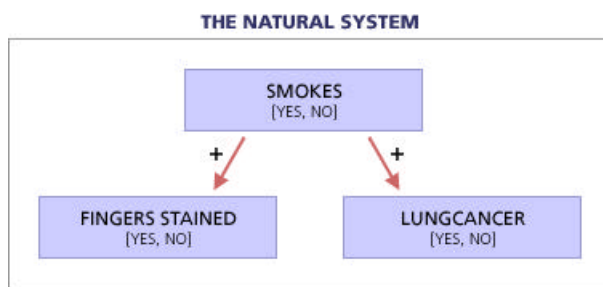


FIGURE 2310-1

The causal graph above describes the natural, or "pre-manipulated system" that produced the data Bert and Russell analyzed prior to suggesting their policy of hand washing. In the natural system, whether a person smokes or not influences whether or not they have stained fingers. In their post-manipulation world, however, everyone is to wash their hands with their special soap every day. What does the causal graph in the "post-manipulation system" involving **FINGER STAINS**, **SMOKING** and **LUNG CANCER** look like?

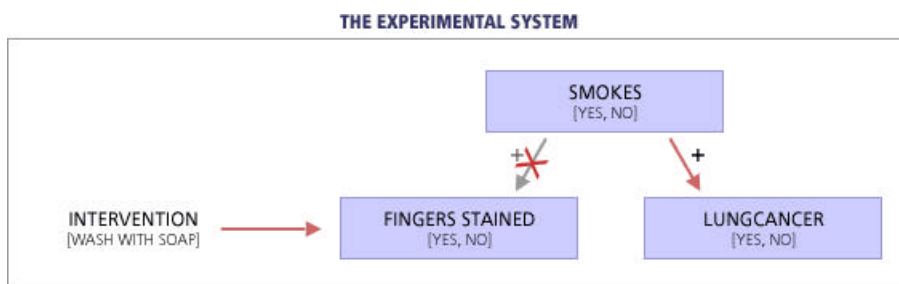


FIGURE 2310-2

In the post-manipulation system, smoking no longer causes finger stains, because everyone is washing their hands with special soap every day. This change in the system eliminates the causal connection in the pre_manipulation system that was producing the association between **FINGER STAINS** and **LUNG CANCER**, so in the post-manipulation setting, **FINGER STAINS** and **LUNG CANCER** won't even be associated.

We can see this experiment as an (imaginary) application of the previous section's lesson. Bert and Russell start out with only an association, which could be explained either by a common cause, or by one of the variables causing the other (or a combination of the two). So, Bert and Russell need to perform an experiment with an ideal intervention. Now, if **FINGER STAINS** actually caused **LUNG CANCER**, then **FINGER STAINS** and **LUNG CANCER** would still be causally connected (after the experiment). If they are still causally connected, then they would be associated. Since they are **not**, in fact, associated after the experiment, then we know that **FINGER STAINS** cannot be a cause of **LUNG CANCER**. Their experiment enabled them to remove some of the ambiguity in the causal story.

2400: Multiple Interventions

Up until now, we have only considered experiments in which we intervened on a single variable. Sometimes this isn't enough, however. Consider again the case of **EDUCATION** and **INCOME**, in which we checked the U.S. Census, and found that **EDUCATION** and **INCOME** are strongly associated. We then hypothetically performed an experiment in which we intervened to set the value of **EDUCATION**. Suppose that we had performed this experiment and determined to our satisfaction that **EDUCATION** in fact is a cause of **INCOME**.

Now suppose that we also measured each person's level of **HAPPINESS** [High, Low]. We could do a second experiment in which we ideally intervene to set **INCOME**, and suppose that we discovered that **INCOME** is a cause of **HAPPINESS**. Since the first experiment told us that **EDUCATION** → **INCOME**, and the second experiment told us **INCOME** → **HAPPINESS**, we know that **EDUCATION** is an indirect cause of **HAPPINESS**.

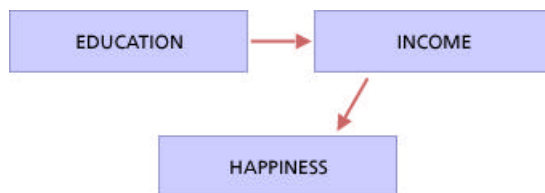


FIGURE 2400-1

How can we tell if **EDUCATION** is also a direct cause of **HAPPINESS**? That is, how can we tell which of the two causal graphs below is accurate?

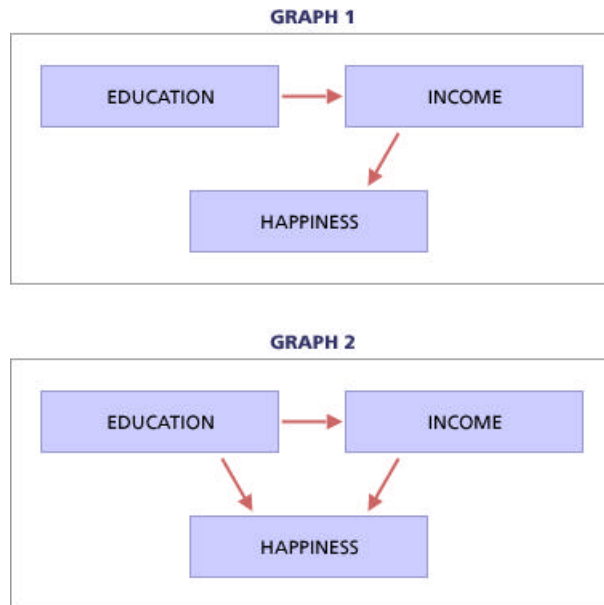


FIGURE 2400-2

Why should this matter? Suppose that you are about to graduate from High School, and although you initially planned to go to college, you are offered an opportunity to make a lot of money right now. You could expect to receive similar job offers upon graduating college, but probably no better than the one you have just been offered. If causal graph 1 is correct, and your main goal is to be happy, then perhaps you should take the job now. Why go for more education if it won't bring extra income, and it won't have any effect on your happiness except through income. If graph 2 is correct, however, then you are better off going to college. It states that acting to get more education will still increase your chances for happiness even if it doesn't increase your income.

< [A link to exercises in the interactive version of this module.](#) >

We can't tell graph 1 from graph 2 by doing an experiment in which we intervene to set **EDUCATION** and asking if **EDUCATION** and **HAPPINESS** are associated, because **both** post-manipulation graphs predict that **EDUCATION** and **HAPPINESS** are associated. If we can experimentally control **INCOME**, however, then we can tell graph 1 from 2.

Consider again our strange experiment to set **EDUCATION**, but this time add the twist that we will experimentally lock **INCOME** at a constant value. In other words, we want to ideally intervene to set the value of **INCOME** to be exactly the same for everyone in the population. Suppose that we ensure that from ages 30 to 35, each of our subjects has an **INCOME** of \$80,000 / year. This intervention is slightly different than the ones we've considered before, since we're going to make sure that everyone has **exactly** the same value. If we now intervene to set each person's education level (though not all to the same value), then **EDUCATION** and **HAPPINESS** will be associated if graph 2 is correct, but they will be independent if graph 1 is correct. Why? Because locking a variable at a fixed value is just one form of an ideal intervention.

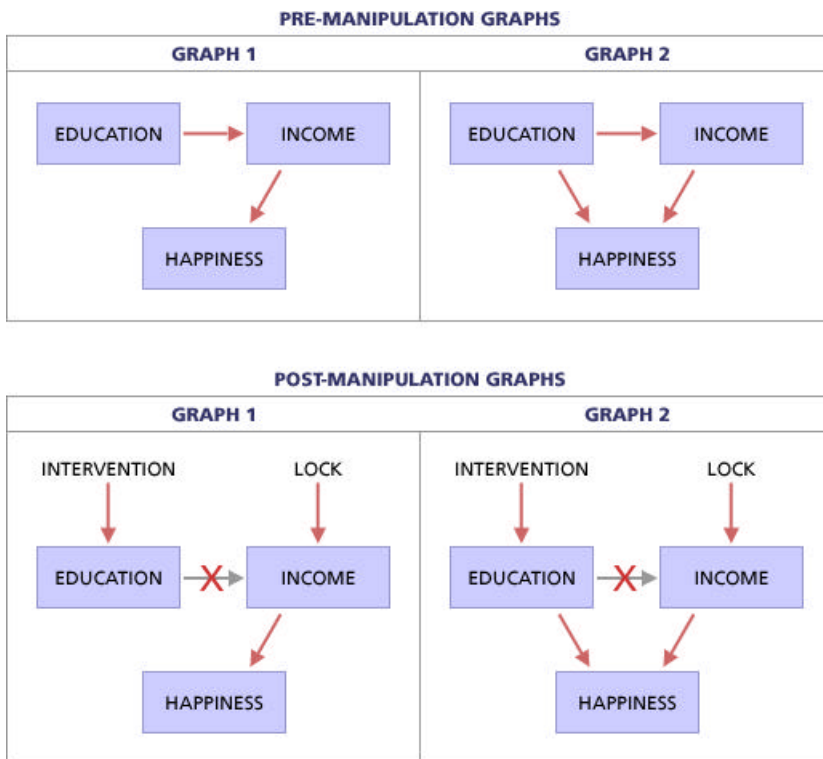


FIGURE 2400-3

In the post_manipulation graphs, you can see that **EDUCATION** and **HAPPINESS** are associated only if **EDUCATION** is a direct cause of **HAPPINESS**, i.e., if graph 2 is the correct theory. So, we can distinguish between the two graphs, but our experiment must have **two** distinct interventions.

3000: Experiments in the Real World

The qualitative theory of how experiments aid causal discovery we just covered is accurate, but a bit idealized. Because ideal interventions are often difficult, and because there are many different ways to intervene, typical experiments are much more complicated than the theoretical analysis suggests. In this section we examine the extra complexity and see why it matters.

Consider an everyday example involving car color and speeding tickets. Suppose it is true that people who own red cars get more speeding tickets than people who own white cars. Suppose you are about to buy a car and you like the color red, so you want to know whether the color of the car you buy will have an effect on whether you get speeding tickets (causal graph 1), or whether the association between car color and tickets is due to some unmeasured confounder, e.g., people who like to speed are also the people who buy red cars (causal graph 2).

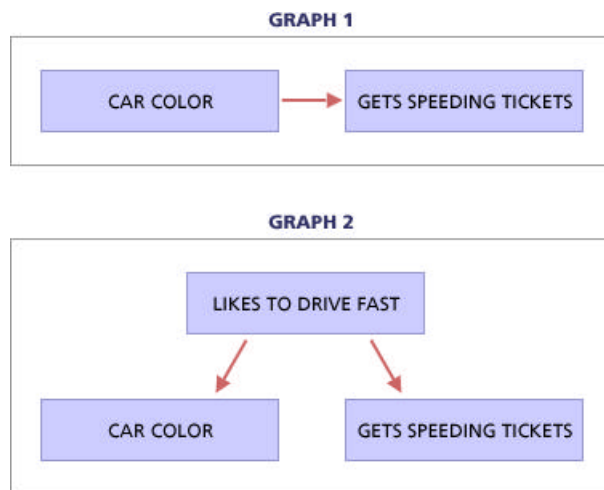


FIGURE 3000-1

If the color affects the chances of getting a ticket, then you will buy a white car, but if the true causal graph is 2, then you might as well buy a red car. How can you find out?

You might perform an experiment in which you intervened to paint lots of cars white that used to be red, and vice versa. You could then observe if the frequency of tickets was higher in the group that started white and went to red, or went down in the group that started red and went to white. Because this is an ideal intervention, it **would** eventually tell us if car color caused speeding tickets. But there is a practical problem looming: because of the problem of uncontrolled causes (explained below), the experiment would take a lot of time and a lot of cars.

How many speeding tickets do most people get in one year? Maybe one or two, but probably none. The average driver gets fewer than one speeding ticket every three years. If we wanted to conduct our experiment in one month, then we would need at least 36 cars before we could expect even a single ticket. Even if we painted 50 red cars white and 50 white cars red, and the red cars got two tickets in the course of the month and the white cars only one ticket, is that enough information to conclude that red cars get more tickets than white cars? No, because its such a small difference that it might be just from luck.

The problem is that lots of things cause getting a speeding tickets -- how fast you are going relative to the speed limit, how fast the regular traffic is going, the road conditions, the mood of the policeman, the time of day, whether it's a holiday or not, whether the car unluckily drove into a speed trap, etc.

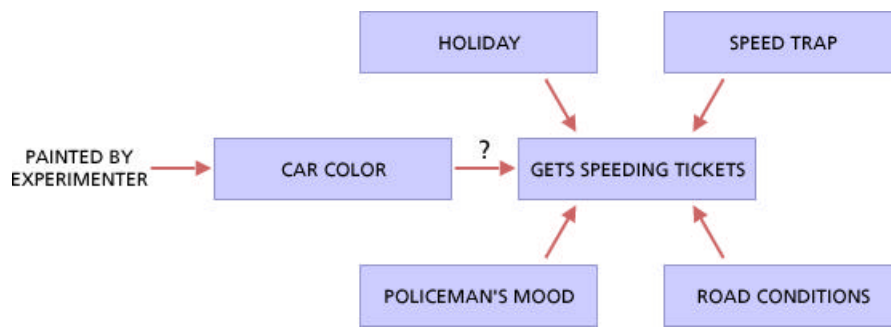


FIGURE 3000-2

Suppose that these other factors vary, which they will, and just by luck in our experiment the red cars happen to confront policemen in bad moods more often than the white cars, then the red cars might get more speeding tickets than the white cars even though car color has no influence. Eventually, if we wait long enough, then the red cars should, on average, confront the same moods in policemen as the white cars. All the other factors besides car color should average out to be the same for the red cars and white cars eventually, but it will take a lot of observations before we can be highly confident that a difference in ticket frequency was due to the color of the cars and not these other factors.

In the experiment we have described, we actually did perform an ideal intervention to set the car's color, but we allowed the other causes of speeding tickets to vary among the individuals in the treatment randomly. As a result, any independence or association that we observe might be due simply to differences in the causes we did not control. So, we can't conclude anything at all about the causal graph, regardless of whether we observe independence or association after the intervention. We will call this the **problem of uncontrolled causes**.

< [A link to exercises in the interactive version of this module.](#) >

4000: Experiments with Matched Groups

4100: The Idea

Our hypothetical one month experiment of painting 50 red cars white and vice versa suffers from the problem of uncontrolled causes. The most appealing alternative is to perform an experiment on car color in which we completely control all the causes, and vary only the color of the car.

Suppose we pick a variety of 10 car models, and then obtain a white and red version of each model that are exactly the same except the color. Now pick a pair of drivers for each pair of cars, and make sure they are the same race, gender, and age, and have them drive the same roads just a few minutes apart, at several different times of day and night. Make sure they both drive at exactly the same speed as each other at all times. Purposely have them drive through speed traps, sometimes at just over the speed limit, sometimes 10 miles an hour over the speed limit, etc. Half the time have the red car go first, and half the time white car go first.

What we have done is set up our experiment so that **all** the causes that might affect getting a ticket are matched in the two groups **except** for car color. The cars are the same make, they are driving at the same speed, on the same roads, and at roughly the same time.

Remember that the **CAR COLOR → GETS SPEEDING TICKET** edge exists if there are values that we can fix the other variables at, such that changing the value of **CAR COLOR** changes the chances of **GETS SPEEDING TICKET**. By matching the causes for the white and red cars, we are checking whether changing just the color changes whether we get a ticket. Hence, if there is any difference in the frequency of tickets between the white cars and their red copies, then the **only** explanation is the difference in car color.

We have intervened on car color -- but we have also experimentally isolated it. We have matched all other factors so that there is no possibility that by luck one group of cars got more tickets from causes other than car color. It's important to note, though, that we did not necessarily intervene on the other causes. We have not intervened to hold fixed the other causes -- we did not determine the mood of the policemen the cars might pass by -- but we have ensured that the other causes have the same values for both groups.

4200: Examples

Lab Mice

When biologists study whether a substance like the red dye in Maraschino cherries causes cancer in mice, they are extremely careful about matching. They don't want the mice who eat the red dye to be different or to be treated differently in any way than the mice who don't eat red dye, because if they were it is possible that by luck the mice who eat the dye might get cancer more frequently because some other factor, and not because of red dye.

So they create pairs of mice that are identical twins -- genetically identical -- and raise them identically except for the red dye. They give them the same diet (except for red dye), the same amount of cage space, the same amount of exercise, etc.

Because they match the twins, the scientists are confident that any difference in the cancer rate between the twins who get the red dye and those who do not can only be from the red dye.

Sneaker Wear

Suppose you are a materials engineer working for Nike, and you discover what you think is a material for shoe soles that will last much longer than the one they are now using. How might you test your new product?

Consider two experiments. In experiment 1, you make 20 pairs of sneakers with soles made of your material (A), and twenty with the current material (B). You randomly select twenty long distance runners to wear sneakers with sole A, and twenty with sole B, and then measure the difference in shoe wear between the two groups.

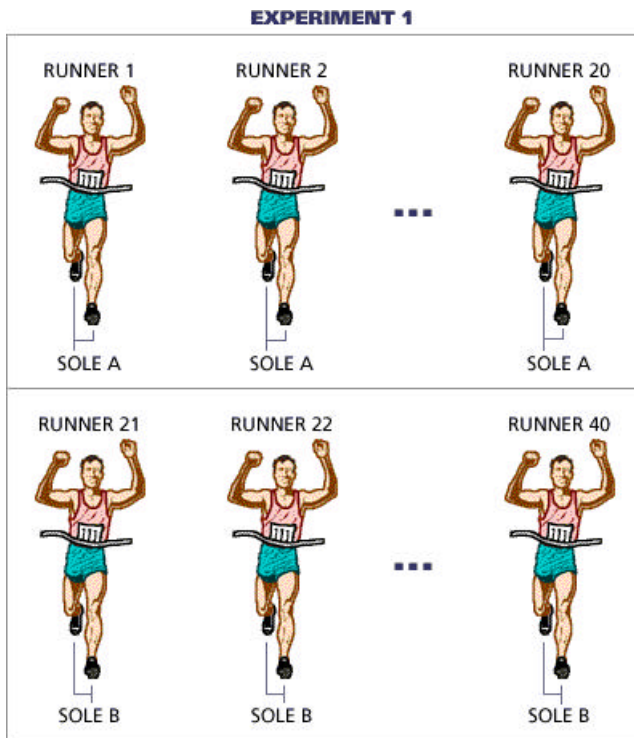


FIGURE 4200-1

In the second experiment, you make twenty pairs of sneakers, in which each pair has one sneaker made with sole material A and one with sole material B. If we assume that a runner puts no more stress on one of his shoes than the other, then all the other causes of shoe wear are the same for each foot except the sole material, unlike in experiment 1.

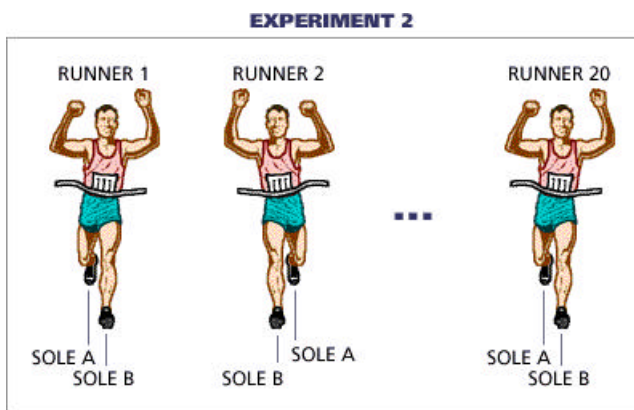


FIGURE 4200-2

< [A link to exercises in the interactive version of this module.](#) >

[5000: Randomization](#)

5100: The Problem

Our ability to control things in an experiment is typically quite limited. In the case of the car color and speeding tickets, and the case of the lab mice, we imagined that we could eliminate all variation between pairs of individuals except the causal factor of interest. Suppose we want to see if a new drug actually controls hay fever. Could we design an experiment in such a way that two individuals varied **only** on whether they took the drug or not? No. Even identical twins, who are genetically identical, cannot be exposed to identical environments from birth until the moment they try the drug under test.

So, we're faced with the following situation: we have a group of individuals (cars, people, etc.), and we need to decide which individuals should receive the treatment (painting the car red, giving the person a new drug, etc.). But the problem is that we need some kind of selection criterion -- a way of assigning each person the treatment or not -- that is not a confounder. In other words, we are worried about the possibility that our selection criteria will be associated with some cause other than the treatment, in which case any observed association between the treatment and effect will be confounded. Before looking at what kind of selection criteria are appropriate in general, let's look at a specific case.

Suppose you built a computer program to tutor statistics students. How would you test whether it was effective? Suppose you wanted to compare how much two individuals learned about statistics, one of whom used your tutor and the other did not. Could you ensure they were identical except for the use of your tutor? No. Since we cannot have matched causes here, we need to use some selection criteria for deciding who will use the tutor and who will not.

We might use the student's height, their last name, their SAT math score, their preference, or any number of other variables. As explained earlier, though, the selection criteria must **not** directly cause the outcome we want to measure. In a graphical representation:

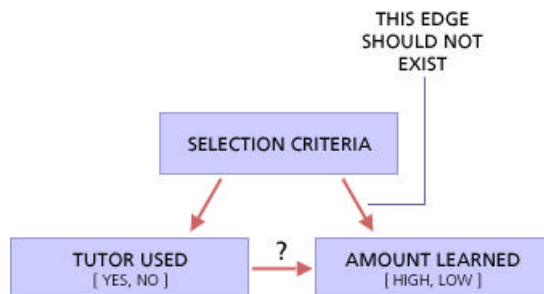


FIGURE 5100-1

If the selection criteria are also a direct cause of **AMOUNT LEARNED**, then we have **introduced** a confounder with our experimental design!

If a chess player means to move his own pawn to threaten his opponent's bishop, but in doing so inadvertently knocks the bishop over as he moves the pawn because his fingers are not sufficiently thin and delicate, we say he has a "Fat Hand." Similarly, if an experiment is set-up such that the selection criterion used for the intervention on a potential cause also directly influences the response, we say that the intervention suffers from being "Fat Hand."

Suppose you decided to conduct an experiment on 200 students, 100 of whom were to use the tutor and 100 not. Consider the following possibilities for selection criteria, and ask whether they would be "Fat Hand."

- 1 Ask for volunteers, and let the first 100 students who volunteer use the tutor.
- 2 Take the 100 students with the highest Math SAT scores and let them use the tutor.
- 3 Assign male students to use the tutor and females to not use it.
- 4 Assign people who sit in the left side of the lecture hall use the tutor, and students who sit in the right side to not use it.

The first two schemes are obvious disasters; they use a variable to intervene on **TUTOR USED** that, in all likelihood, will also directly influence the **AMOUNT LEARNED**. Volunteers are likely to be more eager students and more comfortable with computers, both factors that will directly influence how much the students learn from the course regardless of tutor use. Students with high Math SAT scores will probably learn more from a statistics course because they are comfortable with mathematical ideas and statistics is a mathematical subject. Thus both intervention schemes 1 and 2 are "Fat Hand" interventions that will introduce confounders of **TUTOR USED** and **AMOUNT LEARNED**:

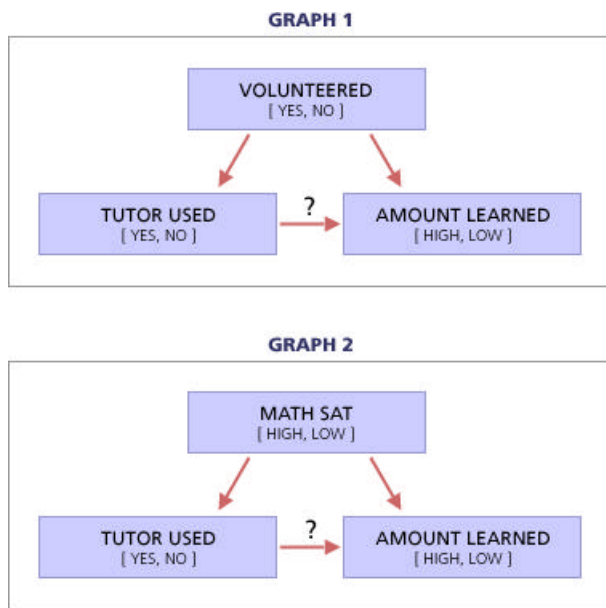


FIGURE 5100-2

What of schemes 3 and 4? Separating tutor use by gender might be OK, unless gender has some effect on the amount learned. Separating tutor use by where you sit is also risky, since the smarter students might sit together.

What we need is a method to select tutor-users that we know is a direct cause of **only** tutor use. In general, if we cannot experimentally match or control **all** the factors that might cause the effect of interest, thus varying only the cause under study between the two groups, then we need a way to select which individuals are assigned the cause of interest so that the selection criterion is independent of all the other causes that might influence the effect. We need an intervention that is ideal (i.e. is not a Fat Hand). This is what Sir Ronald Fisher gave us with Randomized Assignment of Treatment, a technique he introduced in 1935.

5200: Fisher's Solution

Fisher's solution is to use a physical device like a coin or dice and assign treatment based on the outcome of a trial with that device. In the tutor experiment, we might take a fair coin and flip it for each student in the class. If the coin comes up heads, then we assign that student to use the tutor, and if it comes up tails, we assign them to no tutor use.

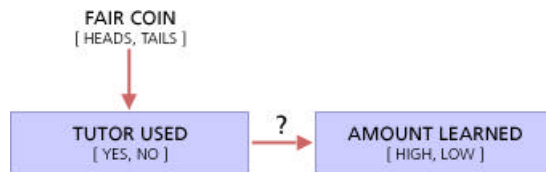


FIGURE 5200-1

Think back to what we wanted from our selection criterion: we needed to guarantee that the selection criterion (or alternately, the property "received the treatment") was not associated with any other causes of the effect. Using a random device clearly satisfies this: in the long run, the frequency of coin flips coming up heads will be 1/2 (assuming a fair coin) regardless of whether we condition on any of the other causes. So in the long run, the assignment of treatment (or not) will be independent of all other causes of the effect in question.

The explanation just given used the phrase "in the long run" several times, which points out an important limitation of randomized assignment of treatment. Although we know that our random device will eventually be independent of all of the other causes, it might not be in the short run, just because of blind luck. If we're unlucky, then after only a few individuals, our random device might happen to be associated with one (or many!) other causes of the effect. In that case, our random device will seem to be a Fat Hand (although that's just because we're unlucky).

For example, even if we flip a coin to assign tutor use, it still might be the case that those students who are assigned to use the tutor happen to have more math aptitude than those who do not, just by bad luck. In that case, we would observe an association that is either entirely or partly due to the association between the selection criterion and math aptitude. In either case, we will incorrectly estimate the causal strength of using the tutor.

At this point, you might ask: how many individuals are enough? When can we say that the association (or independence) we observe is not due to unlucky confounding? Unfortunately, there's no "one size fits all" answer to this question; in fact, we might never have enough individuals (maybe we are just **really** unlucky). However, a major part of statistics is devoted to trying to answer a more fine-grained version of exactly this question: how likely is it that differences we observe in the effect are due to 1) differences in the cause under study, versus 2) differences that arose by chance in the uncontrolled causes?

[6000: Interactive Exploration](#)

[6100: Intervening in the Causality Lab](#)

Using the Randomizer

The Causality Lab gives you the ability to carry out experiments that use randomized assignment of treatment. The Randomizer in the Causality Lab is represented by a die:



FIGURE 6100-1

To use the Randomizer, you simply click on the Randomizer button, and then click on the variable you wish to randomize. When you set up an experiment that uses a randomizer, you are creating an experiment in which you intervene to set the value of the variable to some random value. So, for example, if you place a randomizer on the variable **INCOME** with three values [Low, Medium, and High], then the Causality Lab will automatically set up a random number generator in which, for each individual in your study, they have a 33.33% chance of being assigned each value of **INCOME**.

To become comfortable using the randomizer, first read section 3320 in the Causality Lab User Manual, and then do the following two exercises.

[< A link to exercises in the interactive version of this module. >](#)

Using the Lock

Experiments with randomizers only tell you whether there is a causal path from one variable to another. To find out whether that variable is a **direct** cause, you also need to experimentally lock (i.e., hold fixed) all of the variables that could be mediators on the causal path. The Causality Lab allows you to experimentally hold variables constant using the lock:



FIGURE 6100-2

To use the lock, you first click on the "Lock" button, and then click on the variable you want to lock. You will then have to choose a value for the variable. Once you do, the Causality Lab will automatically ensure that every individual in your study is forced to have that value for the locked variable.

To become comfortable with using the lock, first read section 3330 in the Causality Lab User Manual, and then do the following exercises.

[< A link to exercises in the interactive version of this module. >](#)

6200: Causal Discovery

You now have the core principles of causal discovery with experiments. You know when causal structures produce association and conditional association, you know how experimental interventions work to change pre-manipulation causal graphs into post-manipulation graphs, and how randomizing and experimental control work.

You are thus ready to act like a scientist and do causal discovery with the Causality Lab. If you have not done so, you should now read the Causality Lab User Manual in its entirety.

In this section we give you three exercises in which you must discover the causal structure among a set of variables. Here is a method, slightly vague, but nevertheless a method that you may follow to discover the "true" causal structure:

- 1 First, you should collect data from the pre-manipulated system. Just drag the variables onto the workbench and click on the "Collect a Dataset" button. Make sure to collect "population data" to get around all the problems associated with collecting finite samples.
- 2 Examine the associations and independencies that hold among the variables in the pre-manipulated graph. You can do this by using the "Oracle" in the Independence Tests menu in the Data Display window (where the data is stored).
- 3 Go to the Hypothesize and Make Predictions view in the Workbench. Construct all the causal graphs that you think predict the same independencies and associations that you saw hold in the population data for this particular system. You can check which independencies your hypothesized causal graph entails by clicking on the "Get Independence Facts" button.
- 4 Now perform an experiment (or experiments) of the sort we have been discussing to eliminate one (or more) alternatives. Do this in one of two ways: if you want to know whether one variable is a cause of another, randomize on the first variable, collect "population data," and ask the "Oracle" to tell you what associations hold in the post-manipulation system.
- 5 Alternately, if you are trying to determine whether a variable is a direct or indirect cause, you need to experimentally "lock" the potential causal mediators and compare population data from this experiment to your hypothesized graph's predictions for this experiment.
- 6 Repeat steps 3-5, until you believe you have discovered the structure -- and then click on "Submit".

[< A link to exercises in the interactive version of this module. >](#)

7000: Summary

The scientific search for causal knowledge can be viewed as a struggle against the problem of alternative theories. Faced with limited knowledge and evidence, many causal theories might explain the evidence and be consistent with our background knowledge. We can reduce the number of alternative theories by strengthening our background knowledge; for example, we might assume that variables cannot cause (the values of) variables that were measured earlier. The most powerful technique for eliminating alternative causal theories, though, is an experiment in which we can ideally intervene on a variable.

When we ideally intervene on a variable X , then any variable that was a cause of X in the pre-manipulation system will no longer be a cause of X in the post-manipulation system. So, when we perform an experiment in which we intervene on the variable X , the only causal connections between X and Y that can persist from the pre- to the post-manipulation system are ones in which X is a cause of Y . Since causal connections produce association, if X and Y are associated in the post-manipulation system, then X must be a cause of Y in the pre- and post-manipulation system. Stated again:

No matter how X and Y are causally connected in the pre-manipulation graph, the only causal connection that can hold between them after an ideal intervention on X is a causal path from X to Y .

Thus if X and Y are associated in an experiment in which we ideally intervene to manipulate X , then X is a cause of Y .

There are different ways of actually performing an experiment. When we can fully control or match all the potential causes of Y , then the best possible experimental design compares pairs of individuals that are identical on all causes of Y except for X . In such an experiment, the only explanation for differences in Y is differences in X .

When we cannot fully control or match all of the potential causes of Y , then our experiment suffers from the problem of uncontrolled causes: differences in Y might be due to differences in causes besides X . In addition, we suffer from the danger of assigning values to the cause of interest, X , in a way that introduces a confounder. We suffer from the possibility of using a "Fat Hand" intervention in which the selection criteria for X is also a direct cause of the outcome Y .

The best alternative to complete experimental control or matching is random assignment of treatment, a technique introduced by Fisher in the 1930s. By assigning values to the cause under study from a completely random process like a coin flip, we ensure that our intervention is not a "Fat hand", and that, in the long run, all the other causes of Y , the effect, are independent of the potential cause (that is randomly assigned).
