Introduction

It’s not so hard to come up with a list of factors (independent variables) that might influence biking levels (dependent variable). Designing a study that can prove that any one of these factors actually causes more biking is much trickier. In fact, we can never prove causality definitively in social science and we can never understand causal relationships completely. But if we design studies carefully we can be reasonably confident in concluding that a causal relationship exists. Let’s start with a conceptual model of biking in which living in Davis is positively associated with biking and income is negatively associated with biking.

Questions

1. Finding an empirical association (such as a statistical correlation) between the independent variable and the dependent variable is necessary but not sufficient for establishing causality. Another important criterion in establishing causality is time order: did the variation in the independent variable come before the variation in the dependent variable, or more simply, did the cause come before the effect?

   a. First, how would you use a cross-sectional design to test the association between living in Davis and level of biking for individuals? Hint: How do you make “living in Davis” vary?

   b. Say you use a cross-sectional design and find a positive association between living in Davis and level of biking for individuals. Can you conclude that living in Davis causes more biking? What else could explain this relationship?

   c. What kind of longitudinal design could you use to establish the time order of living in Davis and level of biking?

2. Another important criterion for causality is “nonspuriousness,” that is, that the association between the independent variable and the dependent variable is not due to variation in some third variable extraneous to the conceptual model.

   a. Say you find a negative association between income and level of biking. Can you conclude that having a higher income causes less biking? What other variable(s) that affect both income and biking could explain this relationship?

   b. Say you find a positive association between income and level of biking. Can you conclude that having a higher income causes more biking? What other variable(s) that affect both income and biking could explain this relationship?

   c. What are two different ways you could control for this possibility in your research design?
3. A researcher can feel more confident about a causal relationship if she can explain the causal mechanism, that is, the mechanism by which the independent variable affects the dependent variable. Often, this process involves identifying “mediating variables” (also called “intervening variables”), or the steps along the causal path between the independent variable and the dependent variable.

a. What are some possible mediating variables between living in Davis and level of biking?

b. What are some possible mediating variables between income and level of biking?

4. The last thing to think about with respect to causality is the context in which the causal relationship holds. Say you find good evidence of a causal relationship between income and biking in communities other than Davis. Do you think the causal relationship would still hold in Davis? Why or why not? How might it be different?

5. IF TIME: Choosing an appropriate unit of analysis is also important in explanatory research. So far, we’ve been thinking in terms of individual behavior, with the individual as the unit of analysis. We could also do a study with “city” as the unit of analysis:

a. For this study, what would you use as a dependent variable (think in terms of specific measures)? As independent variables?

b. What would be the unit of observation for the dependent variable?

c. Say that you find that residents of Davis bike more on average than residents of other cities. You also know that Davis residents have higher incomes on average than residents of other cities. It would be tempting to then conclude that income is positively associated with biking at the individual level. Is this necessarily true?

Be sure to turn in your write-up with names of all group members!