correlation between two variables. One explanation is that the $x$ variable does, in fact, cause the $y$ variable. This is known as the causal hypothesis, because it asserts that $x$ (drinking alcohol) causes $y$ (aggressiveness).

In addition to the causal hypothesis, there are two other explanations for why $x$ and $y$ are correlated. One alternative, known as the reverse causation hypothesis, maintains that instead of $x$ causing $y$, $y$ causes $x$. This would occur if a person’s aggressive nature influenced his or her drinking habits. This seems quite plausible. People who are angry and ill-tempered might drink to calm themselves down. A final alternative is known as the third variable problem. This problem occurs when a third variable, usually called $z$, affects $x$ and $y$, but there is no causal relation between $x$ and $y$ in either direction. Stress is a possible third variable in our example. Stress could lead people to drink and to behave aggressively, but the alcohol doesn’t cause aggression and the aggression doesn’t lead to drinking. Other plausible third variables are physical pain, problems at work, or a troubled social life. Conceivably, all of these variables could lead people to drink and to behave aggressively.

It’s important to note that all three processes can occur: $x$ may cause $y$, $y$ may cause $x$, and $z$ may cause both $x$ and $y$. And therein lies the problem. With correlational research, we don’t know which process explains the association between our