OVERVIEW
Most college students can easily complete the phrase, “Correlation does not imply…” with the word “causation.” But how well do they actually apply this reasoning? History is filled with examples of humans’ inherent tendency to infer cause-and-effect from a mere association (e.g., the long-held inference, from correlational data on authoritative parenting and children’s good behavior, that authoritative parenting leads to children’s good behavior). In the abstract, the implication is that when researchers state that “Variable X and Variable Y are correlated, people may mistakenly infer that ‘Variable X causes Variable Y.’”

We hypothesized that the confabulation of correlation with causation is common. We speculated that the order in which variables are presented has an influence on which variable is assumed to be the cause and which the effect; and that the influence of variable order is most robust when correlations are presented both visually and in text. To test these hypotheses, we generated five correlational scenarios and presented each scenario to students in one of six different versions. The versions were created by combining two independent variables: (1) the order of X and Y (before X, before Y, before X and Y); and (2) the form of presentation (text only, scatterplot only, and scatterplot combined). After reading, viewing, or reading and viewing each scenario, participants restated the findings in their own words.

METHOD
A total of 230 students (85M, 145F) across the UWEC campus participated in this study as part of an in-class activity. Students came from a variety of majors; 53% were first-year students, 23% were second-year students, 16% were in their third year, and 10% were in their fourth year or beyond. Each participant responded to five sets of findings. The five scenarios were all presented in the same format for each participant, depending on the condition they had been randomly assigned to. The five sets of findings involved the following pairs of variables:

Women’s self-esteem (X) vs. Women’s number of sex partners (Y)
Women’s IQ (X) vs. Amount of time spent procrastinating (Y)
Children’s level of disobedience (X) vs. Parents’ marital conflict (Y)
Parents’ marital conflict (X) vs. Couples’ emotional intimacy (Y)
Couples’ emotional intimacy (X) vs. Couples’ annual satisfactions (Y)

For half of participants, the variables were presented in a theory order for the first time; the variables were presented in a theory order for the second time. Each of these two groups of participants was split into three subgroups. One group received the findings in a test format only, such as: “Researchers have documented that women’s self-esteem correlates negatively with their number of sex partners. Restate this finding in your own words.” A second group received the findings in the form of a scatter plot only: “Researchers have documented the correlation shown above. Restate this finding in your own words.” A third group received the findings in a combined format: “Researchers have documented that women’s self-esteem correlates negatively with their number of sex partners, as shown above. Restate this finding in your own words.”

RESULTS
Participants’ use of causal language by gender, year in college, and exposure to scientific thinking

Scenario Variable X Variable Y % using causal language Examples of causal language % using non-causal, “fuzzy” language that may reflect causal thinking Examples of non-causal, “fuzzy” language that may reflect causal thinking % of sample who used the same order in their restatements (Chance = 50%) 

ORDER OF VARIABLES IN PARTICIPANTS’ RESTATMENTS
The order in which participants received the variables did not influence their restatements. Instead, the majority of participants restated the findings using the same order as presented in the scenarios. For half of participants, the variables were presented in a theory order for the first time; the variables were presented in a theory order for the second time. Each of these two groups of participants was split into three subgroups. One group received the findings in a test format only, such as: “Researchers have documented that women’s self-esteem correlates negatively with their number of sex partners. Restate this finding in your own words.” A second group received the findings in the form of a scatter plot only: “Researchers have documented the correlation shown above. Restate this finding in your own words.” A third group received the findings in a combined format: “Researchers have documented that women’s self-esteem correlates negatively with their number of sex partners, as shown above. Restate this finding in your own words.”

DISCUSSION
In many psychology classes, instructors demonstrate that correlation does not imply causation through the well known example of the positive association between ice cream sales and murder rate. In this example students have an easy time seeing the absurdity in inferring that ice cream sales have an effect on the murder rate or that the murder rate has an effect on ice cream sales; they easily generate third variables, like warm weather, that explain the association. In most questions that social scientists are interested in, however, the correlations of data are not so easy to recognize. In the current study, we presented students with five correlational scenarios in which either variable could be causal and in which it is possible to generate third variables that could underlie the associations. We found that students did infer causation from data described as correlational, likely because they had preconceived notions about the issues involved. As Stanovich (2010) notes, “when the causal link was made obvious to us, when we have a strong preexisting bias, or when our interpretations become dominated by our aesthetic orientation, it is tempting to treat correlations as evidence of causation.” The current study showed evidence that the majority of students do this even in a circumstance where, if anything, they should have been primed not to (i.e., they were in a classroom, the study was introduced as being about “interpretations of data,” and the scenarios we gave were worded and displayed in a scientific format).

We think the danger of inferring causation from correlation is bigger than some people might think. Take, for example, the scenario we gave participants in which children’s IQ is negatively correlated with the frequency with which they are spanked. Some people might argue that it makes sense to spank children who have low IQs; that spanking precedes children’s IQ, as 70% of our sample did. However, there are several reasons why common sense is not good enough. First, researchers cannot manipulate either IQ or spanking to test their causal effects. Second, it may not occur to people, but it is just as plausible that low IQ can exist without spanking or that a third variable, such as financial dures or parental IQ, could explain both low IQ and spanking. Third, history is full of examples in which the public has invested in mistaken causation, such as the idea that poverty causes schizophrenia and the inference that poverty is causing health problems. Fourth, research in our lab will focus on understanding how people interpret associations between variables under different circumstances. Specifically, we plan to provide specific information about different research designs and ask people to tell us which of a variety of possible conclusions they could infer from the findings.

ACKNOWLEDGMENTS
We thank ORSP for supporting this research and faculty across the UWEC campus for allowing us to collect data in their classrooms: Monty Ernst, John Grauer, Valerie Guient, Alister Iannini, Tracey Jeanet Constable, Carlyne Mirony, and Jeffry Wright.