Background

Sound understanding of research methodology, especially in its applications to causal reasoning, should be considered an essential skill for both students and other individuals outside of the academic community. Society is structured around a constant stream of information; without understanding of the fundamental concepts of research methodology and statistics, one will surely fall behind (Gal, 2002). It is not always easy for individuals to distinguish between causation and correlation (Bleske-Rechek, Morrison, & Heidtke, 2015) and this confusion can have a negative impact on one’s understanding of information they encounter from day to day. Many students have developed their own flawed reasoning, resulting in an absence of this skill (Utts, 2002). For example, is not uncommon to (erroneously) apply the expression “Correlation does not imply causation” when deciding if a relationship between two variables should be considered causal or correlational.

Students develop scientific concepts, or concepts of logic, by internalizing the informational material they receive from their instructors in the classroom (American Psychological Association, 2015). Acquiring new information that contradicts prior knowledge is often a more difficult task compared to when students are completely unfamiliar with the topic (Ambrose, Bridges, Lovett, DiPiero, & Norman, 2010). To make matters worse, a single misconception can ultimately combine with additional misconceptions to craft a complete system of conceptual reasoning (Benassi, Overson, & Hakala, 2014). Despite the level of clarity in which a teacher is able to present information on causal relationships, erasing prior knowledge is not an option; every new piece of information must be related to previously acquired knowledge (Konold, 1995). From this we begin to understand the power behind student misconceptions and
the difficulty associated with the task of teaching skills related to research design to students who might hold similar misconceptions.

In the current study, we aim to understand the specific patterns of error in students’ prior knowledge of research design, particularly with reference to when causal inferences are appropriate. Through manipulations of three key elements of research design (type of treatment, sample size, and number of studies) we can better understand the factors that lead students to make errors in determining when causal inferences are appropriate. Some of these requirements have led to misconceptions regarding strict (often erroneous) guidelines for making causal inferences. We hypothesize that students will be more likely to agree with causal inferences regarding experimental vignettes with a drug/medical treatment compared to a psychological treatment, when sample sizes are larger as opposed to smaller, as when multiple studies are conducted (as contrasted with a single study).

**Methods**

*Participants*

Our first round of data was collected from a sample of 90 undergraduate students (74.4% female) enrolled in six different sections of a pre-service teaching human development course offered at a large Midwestern university. Participants’ ages ranged from 18 to 29 ($M = 20.18$, $SD = 2.07$). Of the respondents who included their ethnicity (78.9%), 90.1% identified as White, 8.5% Black, 1.4% Asian, 1.4% Hispanic, and 1.4% American Indian (*students were able to select more than one race/ethnic group). Of the 90 respondents, 78.9% stated that they were seeking initial teacher certification. Given the relatively small sample size, additional data are currently being collected from the same course during a different semester.
Procedure

The study design was a 2 (type of treatment) × 2 (sample size) × 2 (number of studies) design. Students were randomly presented with one of eight versions of a casual research design prompt during class. Each prompt was constructed with a combination of either drug/medical treatment or psychological treatment (Omega-3 consumption vs. mindfulness training), small or large sample size (40 versus 200 participants), as well as one or multiple studies (one versus five studies). Within the prompt, participants read about a researcher who decides to conduct research on adolescent problem solving ability using either Omega-3 fatty acid consumption or mindfulness training as the treatment. Criteria for a true experimental design were applied to the simulated study, including random assignment to a control or treatment group, homogeneity of demographic characteristics between groups, and equal measures of problem solving ability before the treatment was administered. Thus, correct responses on the prompt would be for students to make both correlational and causal inferences, as correlation is necessary but not sufficient for correlation.

Participants were presented with two five-point Likert scale questions (strongly disagree to strongly agree) where they rated the extent to which they agree or disagree with the following conclusions: 1) the selected treatment caused improvement in problem-solving skills 2) a correlation exists between the selected treatment and problem-solving ability.

Results

Manipulation checks were applied to ensure incorrect responses were the result of misconceptions regarding research design rather than issues with reading comprehension. 19 students who were unable to correctly identify the treatment assigned to the control group, the total number of studies, and the total number of research participants were excluded from the
primary analyses. Two separate factorial ANOVAs with a $2 \times 2 \times 2$ design were conducted to examine the effect of the three manipulations on students’ Likert scale responses for both the causal and correlational questions.

For the causal Likert scale question, preliminary analyses show a significant effect of type of treatment (drug/medical treatment or psychological treatment) on students’ ratings of agreement or disagreement $F(1,63) = 8.61, p < .01$. Contrary to our hypothesis, students who were presented with the drug/medical treatment were less likely to rate the treatment as causing improvement in problem-solving skills ($M=3.23$) compared to students presented with the psychological treatment ($M=3.93$). For the correlational Likert scale question, we found a statistically significant interaction between treatment and sample size $F(1,63) = 4.79, p < .05$. Respondents who were presented with the psychological treatment were more likely to agree with the correlation statement if the sample size was large ($M= 4.48$) than when it was small ($M=3.97$). Also, respondents who were presented with the drug/medical treatment were more likely to agree with the correlation statement if the sample size was small ($M=4.25$) than when it was large ($M=4.00$).

**Discussion**

We will further our study by collecting additional data from students in the same course. The next study will examine students’ inferences about causal and correlational relationships in the context of a correlational research design prompt. We can begin to understand patterns of error in students’ understanding of research design, which results in many of the misconceptions we see regarding causal inferences. With this in mind, instructors of research design courses should be aware that students incorrectly believe the content of treatment to be a factor in whether or not a causal inference can be drawn.
References


