The Journal of General Psychology, 2015, *142*(1), 48–70

Causal Inference from Descriptions of Experimental and Non-Experimental Research: Public Understanding of Correlation-Versus-Causation

APRIL BLESKE-RECHEK
KATELYN M. MORRISON
LUKE D. HEIDTKE
University of Wisconsin-Eau Claire

ABSTRACT. The human tendency to conflate correlation with causation has been lamented by various scientists (Kida, 2006; Stanovich, 2009), and vivid examples of it can be found in both the media and peer-reviewed literature. However, there is little systematic data on the extent to which individuals conflate correlation with causation. In three experiments, we presented people with one of four research vignettes generated from the combination of two independent variables: whether the vignette described an experimental or non-experimental design, and whether it revealed a positive or negative association. Upon reading their vignette, participants selected inferences that could be drawn from the findings. Participants drew causal inferences from non-experimental vignettes as often as they did from experimental vignettes, and more frequently for causal statements and directions of association that fit with intuitive notions than for those that did not. We discuss our findings in relation to other biases in human thinking.

Keywords: causal inference, cognitive biases, correlation and causation, probabilistic trends, thinking

HUMANS ARE COGNITIVE MISERS (STANOVICH, 2009). In modern living, they fall prey to various thinking biases, including the tendency to attend to confirmatory rather than disconfirmatory evidence (Brock & Balloun, 1967), to be influenced by vivid and emotionally compelling cases rather than logic or statistics (Hamill, Wilson, & Nisbett, 1980; Kida, 2006), and to see patterns between independent as well as non-independent events (Hood, 2009; Shermer, 2011). Relatedly, humans have an inherent tendency to look for causes of events and

Address correspondence to April Bleske-Rechek, PhD, University of Wisconsin-Eau Claire, Psychology, 105 Garfield Ave., Eau Claire, WI 54702, USA; bleskeal@uwec.edu (e-mail).

apparent patterns (Hood, 2009; Shermer, 2011) and, in turn, they conflate correlation with causation. This mindware gap (Stanovich, 2009) has been described by authors who study human cognition (Kaplan & Kaplan, 2009) and as noted by Dobelli (2013), the bias "leads us astray practically every day" (p. 110). In the current set of studies we examine systematically the human tendency to conflate correlation with causation.

In psychology, the distinction between correlation and causation is fundamental (APA, 2013). However, the distinction is not unique to psychology. In science more generally, understanding causality and the conditions under which causal inference is appropriate have been long-standing defining features of scientific literacy (Anelli, 2011; Davis, 1935). Indeed, an explicit goal of science education in the United States is to establish high levels of scientific literacy, which involves "understanding the nature of science, the scientific enterprise, and the role of science in society and personal life . . . . Scientific literacy entails being able to read with understanding articles about science in the popular press and to engage in social conversation about the validity of the conclusions . . . . A literate citizen should be able to evaluate the quality of scientific information on the basis of its source and the methods used to generate it" (National Science Education Standards, 1996, pp. 21–22). In essence, then, citizens in the United States are educated with the end goals of understanding the scientific process and being able to evaluate the validity of claims presented in the science they read about in daily life.

Knowledge of the purposes and limitations of different research methods is one element that people use to evaluate the validity of scientific claims. Students trained in the behavioral and social sciences, for example, are taught early on that true experiments (and natural experiments; Rutter, 2007) are the gold standard for making causal claims about behavior (Leary, 2012; Hatfield, Faunce, & Job, 2006), and that observed correlations are merely one step in the path toward causal inference. To infer that Variable X causes Variable Y, that is, one must (1) show that X and Y covary; (2) show that X precedes Y; *and* (3) address alternative explanations for the association. Given some introductory instruction on non-experimental and experimental research designs, most people can easily complete the popular cliché, "Correlation does not imply causation" and apply it to simple cases. For example, they quickly appreciate that a correlation between height and weight does not imply that height causes weight or that weight causes height. They also recognize that a positive correlation between ice cream sales and a murder rate does not imply that ice cream sales lead to an increased murder rate or vice versa, but rather that the correlation is an illusory link explained by confounds such as temperature and population density.

However, for most topics of interest in the social sciences, such as self-esteem and academic performance, the distinction between correlation and causation is not obvious. Perhaps the distinction is difficult for contexts pertaining to human behavior because these contexts tend to be personally relevant, with intuitive causal inferences (Nisbett & Wilson, 1977) that override reflective reasoning. In fact, one path of causal inference often seems to be more intuitive than the other. For example, it has been far more common for people to assume that self-esteem enhances academic achievement than to assume that academic achievement enhances self-esteem (Baumeister, Campbell, Krueger, & Vohs, 2003). Similarly, publicized research on video game playing and aggression has largely focused on video game playing as a cause of aggression rather than a product of aggression (Anderson et al., 2010); and research on pornography use in the context of romantic relationships has focused more on pornography as a cause of relationship dissatisfaction (or disruption) than as a product of relationship dissatisfaction (Gwinn, Lambert, Fincham, & Maner, 2013; Manning, 2006).

In the current research, we sought to examine the degree to which people in the general community draw causal inferences from hypothetical descriptions of experimental and non-experimental research on human behavior. We selected the three topics alluded to above: video game playing and aggression, self-esteem and academic performance, and pornography consumption and relationship satisfaction. We expected that the tendency to inappropriately draw causal inferences from non-experimental research designs would be quite strong—despite the science education standards under which Americans are educated—for at least three reasons. First, we chose topics to which people can easily relate. Armed with personal experiences or vivid testimonials to inform their view, people are prone to intuitive rather than reflective reasoning (Stanovich, 2009; Watts, 2011). Second, we expected that our participants would draw causal inferences from non-causal research designs because media regarding psychological research is notorious for conflating correlation with causation; perhaps the best evidence of it is that faculty in the social sciences use examples from the media (e.g., French, 2012; Kerr, 2008; Newberg, 2011) in their classroom lessons on correlation versus causation. In the three topics we chose to investigate, we found a variety of media headlines that declared causal relationships between variables that had been studied using non-causal methods. For example, a recent headline in the popular magazine *Glamour* indicated that porn can "affect your relationship," despite that the actual article cautioned "correlation doesn't necessarily mean causation" (Kaufman, 2012).

Third, we expected that people in the general community would draw causal inferences from non-causal research designs because evidence suggests that trained social scientists and practitioners conflate correlation with causation. In particular, researchers in teaching and learning have been criticized for using inappropriate causal statements in their articles (Robinson, Levin, Thomas, Pituch, & Vaughn, 2007; Reinhart, Haring, Levin, Patall, & Robinson, 2013); and counseling psychologists have been warned against utilizing counseling practices that have not been supported by randomized experimental designs (Ray et al., 2011). We suspect that these warning calls to educational and counseling psychologists are indicative of the state of research in the social sciences more generally. As noted by Stanovich (2010, p. 150), explicit teaching of strategies for thinking about probabilities and causality is "spotty and inconsistent." Our search for published work on to the three topics of interest in the current research revealed examples of authors conflating correlation with causation. In a search for "video gaming," we found inferences from longitudinal, non-experimental data about the "consequences" of gaming on psychological well-being (Lemmens, Valkenburg, & Peter, 2011); in a search on "self-esteem" and "academic performance," we found inferences from longitudinal, non-experimental data about the "effect" of self-esteem on likelihood of attending college (James & Amato, 2013); and in a search on "pornography" and "relationship satisfaction," we found inferences from cross-sectional, non-experimental data about the "negative impact" of a male partner's problematic pornography use on their female partner's sexual satisfaction (Stewart & Szymanski, 2012).

In summary, we propose that humans are prone to conflate correlation with causation, particularly in the context of personally relevant research pertaining to human behavior. The tendency to falsely infer causality is exacerbated by, perpetuated by, or at least reflected in, both the media and scholarly literature. To study this tendency systematically, we conducted a set of three between-subjects experiments. In each experiment, we exposed people to either an experimental or non-experimental investigation of two variables, and in the investigation the two variables were found to be either positively or negatively associated with each other. By assessing the inferences people draw from the hypothetical research scenarios, we aimed to determine the rate at which people correctly infer causality from experimental research and falsely infer causality from non-experimental research.

Method

Participants

For each study, the authors prepared materials for 120 participants (30 per condition). For a set number of days, they recruited participants from high-traffic restaurants and coffee shops in three different Midwestern communities. Table 1 displays participant demographics. Final sample sizes did not reach 120, for two reasons: first, when scheduled data collection days were over, data collection stopped; second, a variety of individuals agreed to the study and began reading their assigned materials but then either brought them back saying they would have to think too much or they felt the topic was not relevant to them (for example, a widow reading about pornography and men's marital satisfaction). In each study, the sample was approximately 40% male; the typical participant was middle-aged and had some exposure to college.

Materials

Each study was a 2 $\times$ 2 between-subjects experimental design, in which participants read one of four versions of a hypothetical research vignette. In each

TABLE 1. Participant Demographics

Demographic Variable	Study 1 (N = 117)	Study 2 (N = 104)	Study 3 (N = 100)
Participant Sex (% Male)	42%	42%	40%
Participant Age (years)			
Mean	45.72	46.66	39.48
SD	17.86	17.25	16.42
Range	19–80	18–73	19–72
Highest Educational Attainment			
High School Degree or Less	15%	18%	13%
Some College or an Associate's Degree	36%	43%	46%
Bachelor's Degree or More	49%	39%	41%

study the first independent variable was whether the research vignette described an experimental design or a non-experimental design; the second independent variable was whether the research vignette described a positive relationship between the variables of interest or a negative relationship between the variables of interest. After reading the vignette, participants selected the inferences they would draw. On the back side of the questionnaire, participants provided basic demographic information, including their biological sex, age, and education level.

Study 1: Video Games and Aggression

Participants in Study 1 read about a group of teachers who were interested in the association between video games and aggression in their male students. Table 2 provides a layout of the information participants were exposed to in each version of the research vignette. In experimental versions of the vignette, video game playing was described as a manipulation in the teachers' study design. In non-experimental versions, video game playing was described as a measured variable. The experimental and non-experimental vignettes were further split by the direction of the association that the teachers found between video game playing and aggression. In one experimental vignette and one non-experimental vignette, the teachers found a positive (direct) relationship between the number of hours the boys spent playing video games and personnel reports of the boys' aggressiveness; in one experimental vignette and one non-experimental vignette, the teachers found a negative (inverse) relationship between the two variables.

TABLE 2. Study 1: Video Games and Aggression

	Video Game Playing Manipulated		Video Game Playing Measured	
	↑ Video Game ↑ Aggressiveness	↑ Video Game ↓Aggressiveness	↑ Video Game ↑Aggressiveness	↑ Video Game ↓Aggressiveness
Aggressiveness causes an increase in time spent playing video games.	10%	18%	27%	21%
Aggressiveness causes a decrease in time spent playing video games.	3%	4%	7%	10%
Playing video games causes an increase in aggressiveness.	$\boxed{\textbf{63\%}}$	50%	63%	41%
Playing video games causes a decrease in aggressiveness.	10%	$\boxed{\textbf{21\%}}$	3%	17%
Boys who spend more time playing video games tend to be more aggressive.	$\boxed{\textbf{87\%}}$	50%	$\boxed{\textbf{93\%}}$	38%
Boys who spend more time playing video games tend to be less aggressive.	13%	$\boxed{\textbf{36\%}}$	7%	$\boxed{\textbf{41\%}}$

Note. Outlined values indicate the percent of participants who gave the correct response.

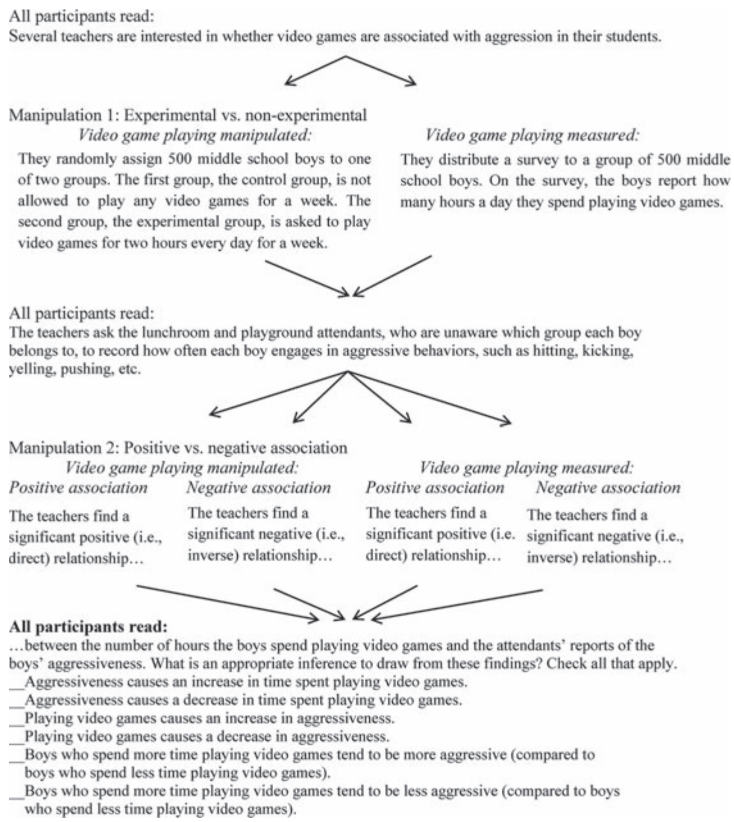

FIGURE 1. Method for Study 1: Video games and aggression.

After reviewing the vignette, participants read six statements and checked which ones would be appropriate inferences for the teachers to make on the basis of their findings. As shown in Figure 1, the first two statements specified aggressiveness as a cause of video game playing in either the positive or negative direction; the next two statements specified video game playing as a cause of aggressiveness in either the positive or negative direction; and the final two statements specified a probabilistic association between the two variables in either the positive or negative direction.

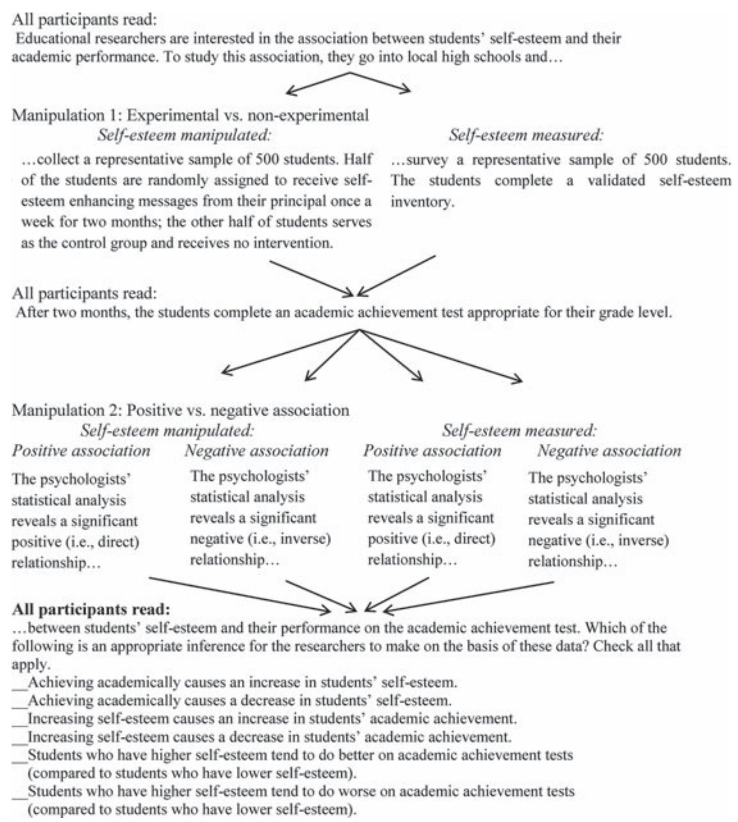

FIGURE 2. Method for Study 2: Self-esteem and academic performance.

Study 2: Self-Esteem and Academic Performance

Participants in Study 2 read about a team of educational researchers who were interested in the association between students' self-esteem and their academic performance. Figure 2 provides a layout of the information participants were exposed to in each version of the research vignette. In experimental versions of the vignette, self-esteem was described as a manipulation in the educational researchers' study design. In non-experimental versions, self-esteem was described as a measured variable. The experimental and non-experimental vignettes were further split by the direction of the association that the researchers found between self-esteem and academic success. In one experimental vignette and one non-experimental vignette, the researchers found a positive (direct) relationship between students' self-esteem and their performance on an academic achievement test; in one experimental vignette and one non-experimental vignette, the researchers found a negative (inverse) relationship between the two variables.

After reviewing the vignette, participants read six statements and checked which ones would be appropriate inferences for the researchers to make on the basis of their findings. As shown in Figure 2, the first two statements specified academic achievement as a cause of self-esteem in either the positive or negative direction; the next two statements specified self-esteem as a cause of academic achievement in either the positive or negative direction; and the final two statements specified a probabilistic association between the two variables in either the positive or negative direction.

Study 3: Pornography Consumption and Marital Satisfaction

Participants in Study 3 read about relationship scientists who were interested in the association between pornography exposure and men's marital satisfaction. Figure 3 a layout of the information participants were exposed to in each version of the research vignette. In experimental versions of the vignette, pornography consumption was described as a manipulation in the relationship scientists' study design. In non-experimental versions, pornography consumption was described as a measured variable. The experimental and non-experimental vignettes were further split by the direction of the association that the scientists found between pornography consumption and marital satisfaction. In one experimental vignette and one non-experimental vignette, the scientists found a positive (direct) relationship between men's pornography consumption and their marital satisfaction; in one experimental vignette and one non-experimental vignette, the scientists found a negative (inverse) relationship between the two variables.

After reviewing the vignette, participants read six statements and checked which ones would be appropriate inferences for the researchers to make on the basis of their findings. As shown in Figure 3, the first two statements specified marital satisfaction as a cause of pornography consumption in either the positive or negative direction; the next two statements specified pornography consumption as a cause of marital satisfaction in either the positive or negative direction; and the final two statements specified a probabilistic association between the two variables in either the positive or negative direction.

Procedure

The procedure for all three studies was identical. Participants reviewed a cover letter; upon their consent to participate, they were randomly assigned to read and respond to one of four versions of a research vignette. Upon completion of their questionnaire, participants were debriefed verbally or in writing, according to their preference.

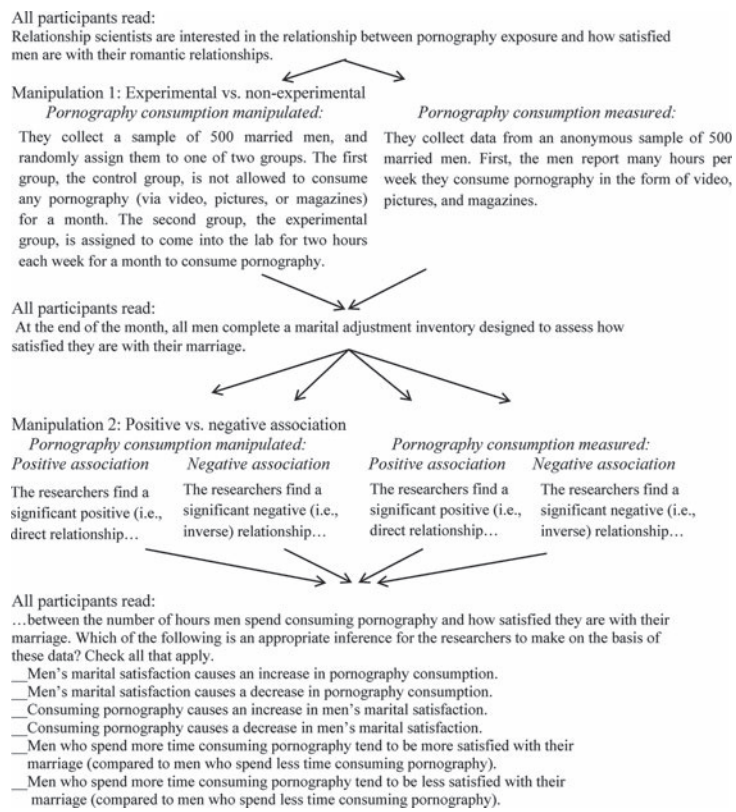

FIGURE 3. Method for Study 3: Pornography consumption and marital satisfaction.

Results

Study 1: Video Games and Aggression

In Study 1, participants read about research pertaining to video game playing and aggressiveness in a sample of boys. Table 2 displays the percent of participants in each condition who selected each inference. Correct inferences are highlighted with boxes.

Causal Inferences from Experimental Vignettes

We first assessed the rate at which participants who were exposed to the experimental vignette (in which video-game playing was manipulated) drew correct causal inferences. Among participants who were told the two variables were found to be positively associated, 63% inferred correctly that playing video games causes an increase in aggressiveness. Among those who were told the two variables were negatively associated, however, only 21% correctly selected the inference that video game playing causes a decrease in aggressiveness. Thus, a greater proportion of participants correctly inferred causation when the influence of video gaming on aggressiveness was presented as positive than when it was presented as negative, $p_a–p_b = .42$, 95% CI [.17, .60], Cramer's $V = .42$. In fact, participants were as likely to incorrectly infer a positive effect of video gaming on aggressiveness after reading that the association was *negative* (50%) as they were to correctly infer a positive effect of video gaming on aggressiveness after reading the association was *positive* (63%), $p_a–p_b = .13$ [–.12, .36], Cramer's $V = .14$.

Causal Inferences from Non-Experimental Vignettes

We next assessed the rate at which participants drew causal inferences from non-causal data. The frequency with which various causal statements were selected by those exposed to the non-experimental vignettes was not reliably different from the frequency with which causal statements were selected by those exposed to the experimental vignettes. For example, among participants who read a non-experimental vignette describing a positive association between video game playing and aggressiveness, 63% incorrectly inferred that video game playing causes an increase in aggressiveness, just as 63% of those who had read the experimental vignette correctly inferred that video game playing causes an increase in aggressiveness, $p_a–p_b = .00$ [−.23, 23], Cramer's $V = .00$. Likewise, among those who read a non-experimental vignette describing a *negative* association, 17% incorrectly inferred that video game playing causes a decrease in aggressiveness, which was similar to the 21% in the experimental condition who had correctly inferred that video game playing causes a decrease in aggressiveness $p_a–p_b = .04$ [−.16, .25], Cramer's $V = .05$.

In non-experimental vignettes, the causal arrow is bidirectional. Video game playing might influence aggressiveness, and aggressiveness might influence video game playing. Theoretically, then, participants should be similarly likely (or unlikely) to infer causality in the two directions. However, participants presented with a non-experimental positive association were more likely to incorrectly conclude that *video game playing* causes an increase in *aggressiveness* (63%) than to incorrectly conclude that *aggressiveness* causes an increase in *video game playing* (27%), OR = 12.0, 95% CI [1.56, 92.29]. Participants presented with a non-experimental negative association were relatively unlikely to make incorrect inferences about the causal effects of aggressiveness and video gaming, with the exception that many (41% [26, 59]) turned the direction of the effect around and incorrectly concluded that video game playing causes an *increase* in aggressiveness.

Understanding the Documented Association Using Different Words

Participants' causal inferences, as described above, suggest that participants hold preconceived notions about how video game playing and aggressiveness go together, and that those notions are biasing their interpretation of the vignettes they read. Further evidence of that bias comes from analysis of our last two statements, which were restatements of the association using slightly different wording. Among those who had read about a positive association, 90% (weighted average of 87% and 93%) correctly selected the positive restatement of that association. However, among those who had read about a negative association, only 39% (weighted average of 36% and 41%) correctly selected the negative restatement of that association. Thus, participants were more likely to select correct restatements of the positive association than of the negative association, $p_a – p_b = .46$ [.30, .59], Cramer's $V = .54$. Moreover, participants exposed to a vignette describing a negative association were just as likely to incorrectly select a statement that described the association as *positive* (44%; weighted average of 50% and 38%) as they were to correctly select a statement that described the association as *negative* (39%; weighted average of 36% and 41%), OR = 1.14 [0.64, 2.02].

Study 2: Self-Esteem and Academic Performance

In Study 2, participants read about research pertaining to students' self-esteem and academic achievement. Table 3 displays the percent of participants in each condition who selected each inference. Correct inferences are highlighted with boxes.

Causal Inferences from Experimental Vignettes

We first assessed the rate at which participants who were exposed to an experimental vignette (in which self-esteem was manipulated) drew correct causal inferences. Among participants who were told the two variables were positively associated, 82% inferred correctly that increasing self-esteem causes an increase in achievement. Among those who were told the two variables were negatively associated, however, only 38% correctly selected the inference that increasing self-esteem causes a decrease in achievement. Thus, participants were more likely to correctly infer causation when the influence of self-esteem on achievement was presented as positive than when it was presented as negative, $p_a – p_b = .44$ [.18, .62], Cramer's $V = .44$. Further, participants presented with a negative causal link were as likely to incorrectly infer that enhancing self-esteem causes an *increase* in achievement (21%) as they were to correctly infer that enhancing self-esteem causes a decrease in achievement (38%), OR = 1.83 [0.68, 4.96].

TABLE 3. Study 2: Self-Esteem and Academic Performance

	Self-esteem Manipulated		Self-esteem Measured	
	↑ Self-esteem ↑ Achievement	↑ Self-esteem ↓ Achievement	↑ Self-esteem ↑Achievement	↑ Self-esteem ↓ Achievement
Achieving academically causes an increase in students' self-esteem.	67%	24%	74%	56%
Achieving academically causes a decrease in students' self-esteem.	4%	0%	0%	32%
Increasing self-esteem causes an increase in students' academic achievement.	**82%**	21%	70%	56%
Increasing self-esteem causes a decrease in students' academic achievement.	7%	**38%**	0%	24%
Students who have higher self-esteem tend to do better on academic achievement tests.	**78%**	28%	**83%**	52%
Students who have higher self-esteem tend to do worse on academic achievement tests.	7%	**38%**	0%	**36%**

Note. Outlined values indicate the percent of participants who gave the correct response.

Causal Inferences from Non-Experimental Vignettes

We next assessed the rate at which participants drew causal inferences from non-causal data. The frequency with which various causal statements were selected by those who were exposed to the non-experimental vignettes was not systematically different from the frequency with which causal statements were selected by those who were exposed to the experimental vignettes. For example, among participants who read a non-experimental vignette describing a positive association between self-esteem and achievement, 70% incorrectly inferred that increasing self-esteem causes an increase in achievement, just as 82% of those who had read the experimental vignette correctly inferred that self-esteem causes an increase in achievement, $p_a–p_b = .12$ [−.12, .35], Cramer's $V = .14$. Among those who read a non-experimental vignette describing a *negative* association, 24% incorrectly inferred that increasing self-esteem causes a decrease in achievement; this was similar to the 38% in the experimental condition who had correctly inferred that increasing self-esteem causes a decrease in achievement, $p_a–p_b = .14$ [−.11, .36], Cramer's $V = .15$.

In non-experimental vignettes, the causal path is bidirectional. Self-esteem might influence achievement, and achievement might influence self-esteem. Theoretically, then, participants should be similarly likely (or unlikely) to infer causality in the two directions. In fact, participants presented with a non-experimental positive association were as likely to incorrectly conclude that self-esteem increases achievement (70%) as they were to incorrectly conclude that achievement increases self-esteem (74%), OR = 2.0 [0.18, 22.06]. Participants presented with a non-experimental *negative* association were as unlikely to infer that self-esteem causes a decrease in achievement (24%) as they were to infer that achievement causes a decrease in self-esteem (32%), OR = 3.0 [0.30, 28.84].

Understanding the Documented Association Using Different Words

Participants' causal inferences, as described above, suggest that participants hold preconceived notions about how self-esteem and achievement go together, and that those notions are biasing their interpretation of the vignettes they read. Further evidence of that bias comes from analysis of our last two statements, which were restatements of the association using slightly different wording. Among those who had read about a positive association, 81% (weighted average of 78% and 83%) correctly selected the positive restatement of that association. However, among those who had read about a negative association, only 37% (weighted average of 38% and 36%) correctly selected the negative restatement of that association. Thus, participants were more likely to correctly re-describe the association between self-esteem and achievement when it was presented as a positive association than when it was presented as a negative association, $p_a–p_b = .45$ [.26, .59], Cramer's $V = .46$. Moreover, participants exposed to a vignette describing a negative association were just as likely to incorrectly select a statement that described the association as *positive* (39%, weighted average of 28% and 52%) as they were to correctly select a statement that described the association as negative (37%, weighted average of 38% and 36%), OR = 1.05 [0.56, 1.97].

Study 3: Pornography Consumption and Marital Satisfaction

Participants in Study 3 read about the association between men's pornography use and marital satisfaction. Table 4 displays the percent of participants in each condition who selected each inference. Correct inferences are highlighted with boxes.

Causal Inferences from Experimental Vignettes

We first assessed the rate at which participants who were exposed to an experimental vignette (the pornography consumption manipulation) drew correct causal inferences. Among participants who were told the two variables were found to be positively associated, 35% inferred correctly that exposure to pornography causes an increase in marital satisfaction. Among those who were told the two variables were negatively associated, however, 69% correctly selected the inference that exposure to pornography causes a decrease in marital satisfaction. In sum, participants were more likely to correctly infer causation when the influence of pornography exposure on men's marital satisfaction was presented as negative than when it was presented as positive, $p_a – p_b = .34$ [.07, .55], Cramer's $V = .34$.

Causal Inferences from Non-Experimental Vignettes

We next assessed the rate at which participants drew causal inferences from non-causal data. The frequency with which various causal statements were selected by those who were exposed to the *non-experimental* vignettes was not systematically different from the frequency with which causal statements were selected by those who were exposed to the *experimental* vignettes. For example, among participants who read a non-experimental vignette describing a positive association between pornography consumption and marital satisfaction, 45% incorrectly inferred that increasing pornography consumption causes an increase in marital satisfaction, just as 35% of those who had read the experimental vignette correctly inferred that increased pornography exposure causes an increase in marital satisfaction, $p_a – p_b = .10$ [–.16, .34], Cramer's $V = .10$. Among those who read a non-experimental vignette describing a *negative* association, 47% incorrectly inferred that pornography consumption causes a decrease in marital satisfaction; this was not reliably different from the 69% in the experimental condition who had correctly inferred that pornography consumption causes a decrease in marital satisfaction, $p_a – p_b = .22$ [–.06, .46], Cramer's $V = .22$.

In non-experimental vignettes, the causal arrow is bidirectional. Pornography consumption might influence marital satisfaction, and marital satisfaction might

TABLE 4. Study 3: Men's Pornography Consumption and Marital Satisfaction

	Pornography Consumption Manipulated		Pornography Consumption Measured	
	↑ Pornography ↑ Satisfaction	↑ Pornography ↓ Satisfaction	↑ Pornography ↑ Satisfaction	↑ Pornography ↓ Satisfaction
Men's marital satisfaction causes an increase in pornography consumption.	22%	14%	28%	21%
Men's marital satisfaction causes a decrease in pornography consumption.	26%	45%	17%	68%
Consuming pornography causes an increase in marital satisfaction.	**35%**	17%	45%	16%
Consuming pornography causes a decrease in marital satisfaction.	13%	**69%**	17%	47%
Men who spend more time consuming pornography tend to be more satisfied with their marriage.	**44%**	10%	**55%**	11%
Men who spend more time consuming pornography tend to be less satisfied with their marriage.	30%	**76%**	21%	**74%**

Note. Outlined values indicate the percent of participants who gave the correct response.

influence pornography consumption. Theoretically, then, participants should be similarly likely (or unlikely) to infer causality in the two directions. Indeed, participants presented with a non-experimental positive association were about as likely to incorrectly conclude that *pornography consumption* causes an increase in *marital satisfaction* (45%) as they were to incorrectly conclude that *marital satisfaction* causes an increase in *pornography consumption* (28%), OR = 2.67 [0.71, 10.05]. Similarly, participants presented with a non-experimental negative association tended to incorrectly infer that marital satisfaction causes a decrease in pornography consumption (68%) about as often as they incorrectly inferred that consuming pornography causes a decrease in marital satisfaction (47%), OR = 3.0 [0.61, 14.86].

Understanding the Documented Association Using Different Words

Participants' causal inferences suggest that they hold preconceived notions about how men's pornography consumption and marital satisfaction are associated. Further evidence of their preconceived notions comes from analysis of our last two statements, which were restatements of the association using slightly different wording. Among those who had read about a positive association, 50% (weighted average of 44% and 55%) correctly selected the positive restatement of that association. However, among those who had read about a negative association, 75% (weighted average of 76% and 74%) correctly selected the negative restatement of that association. Thus, more participants selected a correct restatement of the link between pornography consumption and marital satisfaction when the link was presented as negative than when it was presented as positive, $p_a–p_b = .25$ [.06, .42], Cramer's $V = .26$.

Other Analyses: Inferences by Educational Level

Because participants varied in their level of education, we conducted analyses to determine whether participants with more formal education were more likely than those with less formal education to select correct inferences and to not select incorrect inferences. To maximize cell sizes, we split the participants in each study into two broad education groups: lower (associates degree/some college or less) and higher (bachelor's degree or more). We compared the two groups on three variables of interest for which the cell sizes were large enough to allow for reliable analysis. First, for each study we compared the two education groups, across vignette given, on their likelihood of choosing the correct restatement of what they had read. This analysis included the entire sample. Second, for each study we compared the two education groups on their likelihood, upon reading about a manipulated variable, of choosing its correct corresponding causal inference (in terms of both causal path and direction of association). This analysis involved the set of participants (approximately one-half) who were exposed to a vignette involving a manipulated variable. Third, for each study we compared the two education groups on their likelihood, upon reading about a measured (not manipulated variable), of correctly *not* choosing a causal inference. This analysis involved the set of participants (approximately one-half) who were exposed to a vignette involving a measured variable.

The results of these analyses are displayed in Table 5. First, education level was not consistently associated with participants' likelihood of selecting the correct restatement of the association documented in their specific vignette. In Study 1 regarding video game playing and aggression, participants in the higher education group were more likely than those in the lower education group to choose the correct restatement; but in Study 2 and Study 3 the two education groups did not differ reliably.

Second, education level was not associated with the rate at which those participants who read about a manipulation selected the correct causal inference pertaining to the manipulation. As shown in Table 5, in none of the three studies were those in the higher education group more likely to choose the correct causal inference pertaining to the specific manipulation and direction of outcome they read about.

Third, education level was not consistently associated with the rate at which those participants who read about a non-experimental design responded correctly by not selecting any of the causal inference statements. In Study 2 regarding self-esteem and academic performance, participants in the higher education group were more likely than those in the lower education group to correctly not select any of the causal inferences; but in Study 1 and Study 3 the two education groups did not differ reliably.

Discussion

In this research, we examined the human tendency to conflate correlation with causation. Across the three scenarios, we saw similar patterns of response. First, participants drew causal inferences from non-causal data as often as they did from causal data. For example, participants who read about a study in which video game playing was *measured* and participants who read about a study in which video game playing was *manipulated* were equally likely to infer that video game playing influences aggressiveness. Second, participants were more likely to infer directions of causality that coincided with common-sense notions about the topic. For example, participants who were told that video game playing was associated with aggression were more likely to infer that video game playing caused an increase in aggression, as opposed to a decrease in aggression, regardless of whether the association they were presented with was positive or negative. Similarly, participants who read about self-esteem as a manipulated variable were far more likely to correctly infer its causal influence when it was presented as having a positive effect on achievement than when it was presented as having a negative effect. Notably, correlational research suggests that the actual link

TABLE 5. Percent of Participants With Correct Selections, by Level of Formal Education

	Lower Education	Higher Education	χ^2	p	V
Percent That Selected Correct Restatement of Given Trend					
Study 1: Video Game Playing and Aggression	53%	77%	7.31	.007	.25
Study 2: Self-esteem and Academic Performance	56%	63%	0.40	.529	.06
Study 3: Pornography and Marital Satisfaction	58%	68%	1.17	.280	.11
Percent That Selected Correct Causal Inference (Given Experimental Vignette)					
Study 1: Video Game Playing and Aggression	38%	50%	0.91	.339	.13
Study 2: Self-esteem and Academic Performance	59%	59%	0.00	.984	.00
Study 3: Pornography and Marital Satisfaction	59%	48%	0.60	.438	.11
Percent That Correctly Did Not Select Any Causal Inference (Given Non-experimental Vignette)					
Study 1: Video Game Playing and Aggression	32%	39%	0.28	.599	.07
Study 2: Self-esteem and Academic Performance	3%	28%	6.15	.013	.36
Study 3: Pornography and Marital Satisfaction	20%	28%	0.39	.535	.09

Note. Throughout, $df = 1$. Lower education = associates degree, some college, high-school degree, or less; Higher education = bachelor's degree or equivalent, master's degree or equivalent, doctoral degree or equivalent, or post-doctoral training. Study 1 Lower $N = 60$, Higher $N = 57$; Study 2 Lower $N = 64$, Higher $N = 40$; Study 3 Lower $N = 59$, Higher $N = 41$.

between self-esteem and achievement is weak; and some experimental research suggests that self-esteem enhancements may have a negative effect on achievement (Forsyth, Lawrence, Burnette, & Baumeister, 2007). Third, participants favored certain causal paths over others. For example, participants were more likely to infer that video game playing causes an increase in aggressiveness than they were to infer that aggressiveness causes an increase in video game playing. This pattern is intriguing given that the vignette participants read never indicated the video games were violent. Moreover, systematic research on violent media and aggressiveness supports the causal influence of both variables (Bushman, 1995; Hasan, Begue, Scharkow, & Bushman, 2013; Lemmens & Bushman, 2006). Finally, we found that participants were consistently more likely to choose the correct restatement of the probabilistic association they read about when the direction of the association coincided with their intuitive notions of how the two variables are related. For example, participants who read about a positive association between video game playing and aggressiveness were more willing to choose the correct rephrasing of that association ("Boys who spend more time playing video games tend to be more aggressive than boys who spend less time playing video games than participants,") than were participants who read about a negative association, where far fewer participants chose the correct statement ("Boys who spend more time playing video games tend to be less aggressive than boys who spend less time playing video games."). Although it is possible that participants were as likely to infer an incorrect effect that is more intuitive as to infer a correct effect that is less intuitive even after reading that the less intuitive outcome did occur, it also is possible that some participants did not actually read carefully the unintuitive outcomes because they thought the outcome would be intuitively obvious. If that is the case, those participants would have skipped ahead to the checklist and chosen the statements they assumed would fit based on the initial information given. Unfortunately, we did not incorporate any measures to ensure that participants actually read the true outcomes, and thus we cannot infer clearly the explanation behind participants' tendency to select more intuitive over less intuitive restatements.

Even though participants tended to favor statements that coincided with their preconceived notions, not all participants selected inferences that were intuitive and correct on the basis of the vignette. That is, even when the direction of the causal association fit common-sense beliefs, about a quarter of participants (averaged across studies) still were unwilling to select the correct causal statement. In addition, when the non-causal association fit common-sense beliefs about the variables, about one in five participants was unwilling to select the restatement of the association. One possible explanation is that these participants were suspicious of being tricked and thought they should give a non-intuitive response. Another possibility is that these participants represent the percent of people who do not understand that associations are probabilistic rather than deterministic. For example, they may think that saying "boys who spend more time playing video games tend to be more aggressive" is synonymous to saying "any boy who spends time playing video games will be aggressive." If people do not understand probabilistic trends, then they will perceive an exception to the stated association as refuting it. One direction for future research is to explore whether individuals' understanding of probabilistic trends, assumptions of magic bullet "causes," or reliance on vivid stories of apparent exceptions (Stanovich, 2009) can account for their unwillingness to accept probabilistic statements.

Our findings raise concerns about the tendency of media reports to conflate correlation and causation. If the average reader of research findings is prone to inferring causation from correlation, and does not have explicit training in the inferences allowed from various research designs, then those delivering the news need to do their best to not mislead the already susceptible reader. Media headlines and titles seem to frequently advertise the "effects" of one variable on another despite that the research under discussion does not permit causal claims. Causal claims about correlational data came from a wide variety of domains, such as spanking and aggression (Boyles, 2005), sharing chores and marital satisfaction ("Sharing chores," 2013), binge drinking and heart problems (Jaslow, 2013), childhood candy consumption and violent offense arrests (Martinez, 2009), and texting and language skill ("Texting improves," 2009). One direction for future research is to document how frequently media sources use headlines or article texts that inappropriately utilize causal language in the context of correlational research.

In conclusion, we documented that people in the community frequently draw causal inferences from non-causal data, and they draw inferences—in terms of direction and causal path—that fit with their intuitive notions, regardless of the findings presented. Although our samples were relatively homogeneous in ethnic diversity, they did vary widely in age and educational background, and the pattern of responses was consistent across the three scenario studies. Moreover, participants of varying levels of education were similarly likely to err in their inferences. We hope that future research in this area will focus on further understanding why people so readily conflate correlation with causation, the potential impact of media on people's understanding of probabilistic trends, and how scientists and educators can intervene to enhance people's ability to reason about causal and non-causal trends.

AUTHOR NOTES

April Bleske-Rechek is a psychology professor at the University of Wisconsin-Eau Claire. Her teaching and scholarship interests involve research methods and science-mindedness, individual differences, and evolutionary psychology. **Katelyn M. Morrison** is an advanced undergraduate student of psychology and health administration at the University of Wisconsin-Eau Claire. **Luke D. Heidtke** earned his BA at the University of Wisconsin-Eau Claire. He is now pursuing his Master's degree in counseling at the Univeristy of Wisconsin-Stout.

REFERENCES

American Psychological Association. (2013). *APA guidelines for the undergraduate psychology major, Version 2.0*. Washington, DC: APA.

Anderson, C. A., Shibuya, A., Ihori, N., Swing, E. L., Bushman, B. J., Sakamoto, A., & Saleem, M. (2010). Violent video game effects on aggression, empathy, and prosocial behavior in Eastern and Western countries: A meta-analytic review. *Psychological Bulletin*, *136*, 151–173.

Anelli, C. (2011). Scientific literacy: What is it, are we teaching it, and does it matter? *American Entomologist*, *57*, 235–243.

Baumeister, R. F., Campbell, J. D., Krueger, J. I., & Vohs, K. D. (2003). Does high self-esteem cause better performance, interpersonal success, happiness, or healthier lifestyles? *Psychological Science in the Public Interest*, *4*, 1–44.

Boyles, S. (2005, November 14). Can spanking cause aggression? *CBS News*. Retrieved from http://www.cbsnews.com/news/can-spanking-cause-aggression/

Brock, T. C., & Balloun, J. L. (1967). Behavioral receptivity to dissonant information. *Journal of Personality and Social Psychology*, *6*, 413–428.

Bushman, B. J. (1995). Moderating role of trait aggressiveness in the effects of violent media on aggression. *Journal of Personality and Social Psychology*, *69*, 950–960.

Davis, I. C. (1935). The measurement of scientific attitudes. *Science Education*, *19*, 117–122.

Dobelli, R. (2013). *The art of thinking clearly*. New York, NY: Harper Collins.

Forsyth, D. R., Lawrence, N. K., Burnette, J. L., & Baumeister, R. R. (2007). Attempting to improve the academic performance of struggling college students by bolstering their self-esteem: An intervention that backfired. *Journal of Social and Clinical Psychology*, *26*, 447–459.

French, C. (2012, February 7). Spanking kids can cause long-term harm-Canada study. *Reuters*. Retrieved from http://www.reuters.com/article/2012/02/07/canada-spanking-idUSL2E8D1F1C20120207

Gwinn, A. M., Lambert, N. M., Fincham, F. D., & Maner, J. K. (2013). Pornography, relationship alternatives, and intimate extradyadic behavior. *Social Psychological and Personality Science*, *4*, 699–704.

Hamill, R., Wilson, T. D., & Nisbett, R. E. (1980). Insensitivity to sample bias: Generalizing from atypical cases. *Journal of Personality and Social Psychology*, *39*, 578–589.

Hasan, Y., Begue, L., Scharkow, M., & Bushman, B. J. (2013). The more you play, the more aggressive you become: A long-term experimental study of cumulative violent video game effects on hostile expectations and aggressive behavior. *Journal of Experimental Social Psychology*, *49*, 224–227.

Hatfield, J., Faunce, G. J., & Job, R. F. (2006). Avoiding confusion surrounding the phrase 'correlation does not imply causation.' *Teaching of Psychology*, *33*, 49–51.

Hood, B. M. (2009). *The science of superstition: How the developing brain creates supernatural beliefs*. New York, NY: Harper Collins.

James, S. L., & Amato, P. R. (2013). Self-esteem and the reproduction of social class. *Social Science Quarterly*, *94*, 933–955.

Jaslow, R. (2013, April 25). Binge drinking in college may lead to heart problems later in life. *CBS News.* Retrieved from http://www.cbsnews.com/news/binge-drinking-in-college-may-lead-to-heart-problems-later-in-life/

Kaplan, M., & Kaplan, E. (2009). *Bozo Sapiens: Why to err is human*. New York, NY: Bloomsbury Press.

Kaufman, G. (2012, June). How much does your guy like porn? Here's how it might affect your relationship . . . *Glamour.* Retrieved from http://www.glamour.com/sex-love-life/blogs/smitten/2012/06/how-much-does-your-guy-like-po.html

Kerr, J. C. (2008, May 13). Feds: Teen use of pot can lead to dependency, mental illness. *The Seattle Times*. Retrieved from http://seattletimes.com/html/health/2004402313_apteensdrugs.html

Kida, T. E. (2006). *Don't believe everything you think: The 6 basic mistakes we make in thinking*. Amherst, NY: Prometheus Books.

Leary, M. R. (2012). *Introduction to behavioral research methods* (6th ed.). Boston, MA: Pearson.

Lemmens, J. S., & Bushman, B. J. (2006). The appeal of violent video games to lower educated aggressive adolescent boys from two countries. *CyberPsychology & Behavior*, *9*, 638–641.

Lemmens, J. S., Valkenburg, P. M., & Peter, J. (2011). The effects of pathological gaming on aggressive behavior. *Journal of Youth and Adolescence*, *40*, 38–47.

Manning, J. C. (2006). The impact of Internet pornography on marriage and the family: A review of the research. *Sexual Addiction & Compulsivity*, *13*, 131–165.

Martinez, E. (2009, October 1). Sugar rush . . . to prison? Study says lots of candy could lead to violence. *CBS News*. Retrieved from http://www.cbsnews.com/news/sugar-rush-to-prison-study-says-lots-of-candy-could-lead-to-violence/

National Science Education Standards. (1996). *National Research Council's National Committee on Science Education Standards and Assessment*. Washington, DC: National Academies Press.

Newberg, A. (2011, May 31). Religious experiences shrink part of the brain. *Scientific American*. Retrieved from http://www.scientificamerican.com/article/religious-experiences-shrink-part-of-brain/

Nisbett, R E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*, 231–259.

Ray, D. C., Hull, D. M., Thacker, A. J., Pace, L. S., Swan, K. L., Carlson, S. E., & Sullivan, J. M. (2011). Research in counseling: A 10-year review to inform practice. *Journal of Counseling & Development*, *89*, 349–359.

Reinhart, A. L., Haring S. H., Levin, J. R., Patall, E. A., & Robinson, D. H. (2013). Models of not-so-good behavior: Yet another way to squeeze causality and recommendations for practice out of correlational data. *Journal of Educational Psychology*, *105*, 241–247.

Robinson, D. H., Levin, J. R., Thomas, G. D., Pituch, K. A., & Vaugh, S. (2007). The incidence of 'causal' statements in teaching-and-learning research journals. *American Educational Research Journal*, *44*, 400–413.

Rutter, M. (2007). Proceeding from observed correlation to causal inference. *Perspectives on Psychological Science*, *2*, 377–395.

Sharing chores leads to a happy marriage. (2013, April 25). *Express*. Retrieved from http://www.express.co.uk/news/uk/394527/Sharing-chores-leads-to-a-happy-marriage

Shermer, M. (2011). *The believing brain*. New York, NY: Times Books.

Stanovich, K. E. (2009). *What intelligence tests miss: The psychology of rational thought*. New Haven, CT: Yale University Press.

Stanovich, K. E. (2010). *How to think straight about psychology* (9th ed.). Boston, MA: Allyn & Bacon.

Stewart, D. N., & Szymanski, D. M. (2012). Young adult women's reports of their male romantic partner's pornography use as a correlate of their psychological distress, relationship quality, and sexual satisfaction. *Sex Roles*, *67*, 257–271.

Texting 'improves language skill.' (2009, February 25). *BBC*. Retrieved from http://news.bbc. co.uk/2/hi/technology/7910075.stm

Watts, D. J. (2011). *Everything is obvious: Once you know the answer*. New York, NY: Random House.

Original manuscript received June 4, 2014
Final version accepted October 13, 2014