

# Statistical Control Requires Causal Justification



Anna C. Wysocki<sup>1</sup>, Katherine M. Lawson<sup>1</sup>, and Mijke Rhemtulla<sup>1</sup>

Department of Psychology, University of California, Davis

Advances in Methods and  
 Practices in Psychological Science  
 April-June 2022, Vol. 5, No. 2,  
 pp. 1–19  
 © The Author(s) 2022  
 Article reuse guidelines:  
 sagepub.com/journals-permissions  
 DOI: 10.1177/25152459221095823  
 www.psychologicalscience.org/AMPPS



## Abstract

It is common practice in correlational or quasiexperimental studies to use statistical control to remove confounding effects from a regression coefficient. Controlling for relevant confounders can debias the estimated causal effect of a predictor on an outcome; that is, it can bring the estimated regression coefficient closer to the value of the true causal effect. But statistical control works only under ideal circumstances. When the selected control variables are inappropriate, controlling can result in estimates that are more biased than uncontrolled estimates. Despite the ubiquity of statistical control in published regression analyses and the consequences of controlling for inappropriate third variables, the selection of control variables is rarely explicitly justified in print. We argue that to carefully select appropriate control variables, researchers must propose and defend a causal structure that includes the outcome, predictors, and plausible confounders. We underscore the importance of causality when selecting control variables by demonstrating how regression coefficients are affected by controlling for appropriate and inappropriate variables. Finally, we provide practical recommendations for applied researchers who wish to use statistical control.

## Keywords

statistical control, causality, confounder, mediator, collider, open materials

Received 10/10/20; Revision accepted 3/19/22

Psychological research that uses observational or quasiexperimental designs can benefit from statistical control to remove the effect of third variables—variables other than the target predictor and outcome—from an estimate of the causal effect that would otherwise be confounded (Breugh, 2006; McNamee, 2003).<sup>1</sup> Statistical control can lead to more accurate estimates of a causal effect (Pearl, 2009), but only when the right variables are controlled for (Rohrer, 2018).<sup>2</sup> Although controlling for third variables is common practice (Atinc et al., 2012; Bernerth & Aguinis, 2016; Breugh, 2008), the selection of these variables is rarely justified on causal grounds.

In this article, we illustrate that controlling for an inappropriate variable can result in biased causal estimates. We begin with a brief introduction to causal inference and regression models, a definition of statistical control, and a description of situations in which statistical control is useful for researchers. We then highlight the pervasive issues that surround how control variables are typically selected in psychology. We outline the assumptions required to justify controlling for a third

variable—most importantly, that the control variable is a plausible confounder or lies on the confounding path. Next, we discuss the consequences of controlling for other types of third variables, including mediators, colliders, and proxies. We then discuss how longitudinal data can be used to deal with more complex models. Using an applied example, we provide practical recommendations for applied researchers who work with observational data who wish to use statistical control to bolster their causal interpretations.

## A Brief Introduction to Causal Inference

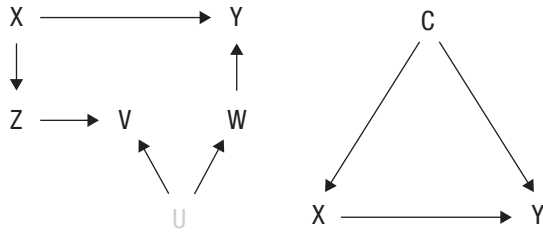
In this section, we provide a very brief introduction to the field of causal inference, including key concepts and definitions that are relevant for the present article. The field of causal inference expands far beyond what we

### Corresponding Author:

Anna C. Wysocki, Department of Psychology, University of California, Davis  
 Email: awysocki@ucdavis.edu



Creative Commons NonCommercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits noncommercial use, reproduction, and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).



**Fig. 1.** Directed acyclic graphs (DAGs). (Left) An example DAG with one unmeasured variable,  $U$ . (Right) A simple example of confounding.

can cover here. We direct interested readers to Dablander (2020) for a slightly lengthier introduction and to Pearl et al. (2016) or Peters et al. (2017) for book-length introductions.

Causal inference involves estimating the magnitude of causal effects given an assumed causal structure. We use Pearl’s (1995) definition of causality, namely, that  $X$  is a cause of  $Y$  when an intervention on  $X$  (e.g., setting  $X$  to a particular value) produces a change in  $Y$ . A causal effect—the expected increase in  $Y$  for a 1-unit intervention in  $X$ —is identified when it is possible to derive an unbiased estimate of the causal effect from data. Estimating the magnitude of causal effects is key to understanding psychological phenomena; however, causal inference relies on theoretical assumptions that come from prior knowledge in addition to statistical information. Identifying a single causal effect requires accounting for and removing confounding effects without inducing spurious effects. Thus, although it is common practice in psychological research to estimate and interpret multiple coefficients simultaneously (e.g., using a regression model with six predictors), we focus on the identification of a single causal effect at a time using a tool called a directed acyclic graph (DAG).

A DAG depicts hypothesized causal relations between variables (for a detailed introduction to DAGs in psychology, see Rohrer, 2018). The DAGs used throughout this article contain three features: capital letters that represent variables; the letter  $U$ , which represents a set of unmeasured variables; and arrows that represent causal effects. For example, Figure 1 (left) depicts that  $X$  is a cause of  $Y$ ,  $Z$ , and, indirectly,  $V$  ( $X$  affects  $V$  through the mediating variable  $Z$ ), and there is an unmeasured common cause that affects both  $V$  and  $W$ .<sup>3</sup> We use DAGs to represent causal relations in the population, and in the same way, researchers can use DAGs to encode hypothesized causal systems and to inform decisions about which statistical models enable the identification of a causal effect. Note that DAGs are nonparametric; that is, they do not require any particular functional form. However, throughout this article, we assume that all causal effects are linear and that each variable has a normally distributed residual.

In a DAG, a *path* is a sequence of arrows that connects one variable to another. A path that contains two

variables will transmit association (i.e., the path results in the two variables being associated) unless there is an *inverted fork* (i.e., two arrowheads coming together; e.g.,  $X \rightarrow Y \leftarrow W$ ) anywhere along the path. Each pair of variables may be connected by multiple paths, and if one of these paths transmits an association, then the pair of variables is expected to be associated.<sup>4</sup> In Figure 1 (left),  $X$  and  $Y$  are connected by two paths:  $X \rightarrow Y$  and  $X \rightarrow Z \rightarrow V \leftarrow U \rightarrow W \rightarrow Y$ . The latter path does not transmit association because of the inverted fork,  $Z \rightarrow V \leftarrow U$ . But the path  $X \rightarrow Y$  does transmit association, and thus, we expect  $X$  and  $Y$  to be associated in sample data drawn from a population represented by this causal graph.

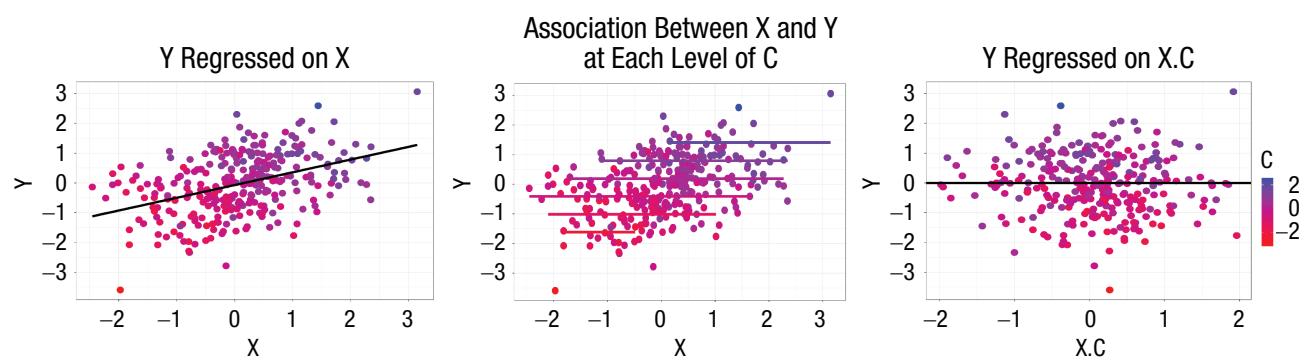
The association information embedded in DAGs can help researchers discover potential threats to identifying a causal effect. We use “bias” to mean the discrepancy between a population parameter (e.g., a population regression coefficient) and the causal effect. In Figure 1 (right), there are two paths that connect  $X$  and  $Y$  and transmit association—both paths contribute to the association between  $X$  and  $Y$ . Crucially, one path,  $X \leftarrow C \rightarrow Y$ , is *noncausal*, that is, it is not part of the causal effect of  $X$  on  $Y$ , so manipulating  $X$  does not change  $C$  or  $Y$  through  $C$ . Thus, the association between  $X$  and  $Y$  is a biased estimate of the causal effect. In general, a common cause (i.e., a confounder) of a predictor and an outcome results in an association that is biased for the causal effect. To remove this bias, the common cause path (in this case, the path through  $C$ ) must be removed from the estimated association. This can be done through an experimental research design in which the predictor is randomized (Greenland, 1990), but experimental manipulation of psychological variables is often unfeasible. Thus, psychologists have had to find another method to block confounding paths. One such method is statistical control, which can be accomplished via regression (McNamee, 2005).

## Linear Regression and Statistical Control

In this section, we review how multiple linear regression produces coefficients that represent the linear association between each predictor and the outcome variable, conditional on the set of other predictors. To simplify the presentation, we assume all variables are standardized (means = 0, variances = 1). The linear regression model formulates an outcome variable,  $Y$ , as a linear function of a set of  $p$  predictor variables,  $X_1, \dots, X_p$ , plus a normally distributed residual with mean 0:

$$Y = \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon.$$

In simple regression ( $p = 1$ ),  $\beta_1$  is equivalent to the correlation between  $X_1$  (which for simplicity we denote



**Fig. 2.** Visualizing statistical control. (Left) The confounded regression of  $Y$  on  $X$  reveals a strong linear association. The color of individual points represents their values on the confounding variable,  $C$ . (Middle) Within each level of  $C$ , the association between  $X$  and  $Y$  is 0. (Right) When  $C$  is included as an additional predictor in the regression equation of  $Y$  on  $X$ , its confounding influence is removed to reveal only the partial association between  $X$  and  $Y$ .  $X.C$  is the residual scores that are obtained when  $X$  is regressed on  $C$ .

as  $X$  in the  $p = 1$  case) and  $Y$  (when all variables are standardized). Likewise, when  $p > 1$  and all predictors are perfectly uncorrelated with each other, then the  $\beta$  for each predictor is equal to the correlation between that predictor and the outcome. In these special cases, each  $\beta$  reflects the total linear association between the predictor and  $Y$  and can be interpreted as the expected change in  $Y$  for a 1-unit change in  $X$ .

When the predictors are correlated with each other, the estimation and interpretation of regression coefficients are more complicated. In these cases, the variance shared among predictors gets partitioned among the regression coefficients such that each regression coefficient represents the expected change in  $Y$  for a 1-unit increase in  $X$ , holding all other predictors fixed. Conceptually, this approach is the statistical equivalent of sampling participants who have the same value on all but one of the predictors and estimating the association between that one predictor and  $Y$  in the sample. Thus, multiple regression coefficients are known as *partial* regression coefficients because they represent the isolated association between a single  $X$  and  $Y$  when none of the other predictors are changing. If the predictors other than  $X$  represent the full set of variables that confounds the  $X$ - $Y$  association and if there is no reverse causality (i.e.,  $Y$  does not cause  $X$ ), then the causal effect ( $X \rightarrow Y$ ) is identified using the partial regression coefficient.

Figure 2 illustrates statistical control. The scatterplot on the left shows a strong linear association between  $X$  and  $Y$ , and the color of the points in this plot represents values of a third variable,  $C$ .  $C$  is correlated with both  $X$  and  $Y$ , which raises the possibility that it may be a confounder. The middle plot shows the population-level regression lines that one would get if it were possible to compute the simple regression coefficient of  $Y$  on  $X$  for each subpopulation with a fixed value on  $C$ . In practice, however, there is not enough information in this

small data set to accurately estimate the regression coefficient within each subpopulation sample. The plot on the right shows the statistical approach to do this, that is, the deconfounded association between  $X$  and  $Y$ . The  $x$ -axis now represents the residuals that are obtained when  $X$  is regressed on  $C$ , that is, the part of  $X$  that is independent of  $C$  ( $X = \beta_1 C + X.C$ ). When  $Y$  is regressed on this residualized predictor,  $Y = \beta_2 X.C + \epsilon$ , there is little remaining relation between  $X.C$  and  $Y$ , and  $\beta_2$  is close to 0. Adding the control variable  $C$  to the regression of  $Y$  on  $X$ ,  $Y = \beta_2 X + \beta_3 C + \epsilon$ , is equivalent to regressing  $Y$  on the residualized  $X$ —the value of  $\beta_2$  is the same in both equations.

By statistically controlling for the correct variable, a confounding effect can be removed from an estimate, which makes statistical control a valuable tool for researchers who are interested in causal inference and have access to observational or quasiexperimental data (Morabia, 2011; Pourhoseingholi et al., 2012). Here, we focus on controlling for the correct variable or variables, but obtaining an unbiased association depends on several additional assumptions: (a) Any interactions or non-linear effects must be specified correctly (Cui et al., 2009; Simonsohn, 2019), (b) the predictor and control variables must be measured without error or a model that deals with the measurement error must be used (Savalei, 2019; Westfall & Yarkoni, 2016), and (c) the relevant variables must be measured at a time when the causal process can be captured.<sup>5</sup>

Causal inference is not the only reason that a researcher may choose to control for a third variable. Controlling for a variable that shares variance with the outcome but not the predictor will decrease the amount of residual variance in the outcome, which, in turn, lowers the standard error of the estimated regression coefficient and increases power (Cohen et al., 2003). Control variables are also sometimes used to establish that a new measure is uniquely predictive beyond some already

established measure (Wang & Eastwick, 2020; we discuss this practice later in the paper). Researchers may also be interested in using the partial regression coefficients to describe partial associations rather than using them to explain psychological processes.

Frequently, however, researchers aim to develop and test theories of psychological processes, and this endeavor almost always involves making and testing causal hypotheses. Although it is rare that causal inference is explicitly acknowledged as the goal in nonexperimental studies (Grosz et al., 2020), a key component of theory building is proposing a set of principles that explains a process and then formulating a model based on these principles (Borsboom et al., 2021)—in other words, positing a set of causal hypotheses (Shmueli, 2010; Yarkoni & Westfall, 2017). When coefficients are interpreted as reflecting the strength of a causal effect, then selecting and controlling for the correct variable is of the utmost importance because controlling for the wrong variable can increase rather than decrease bias.

### Common Practices: How Do Researchers Typically Choose Control Variables?

Reviews of published studies in psychology journals have found that more than 50% of the studies reviewed gave no justification for the inclusion of specific control variables (Becker, 2005; Bernerth & Aguinis, 2016; Breugh, 2008). Moreover, Atinc and colleagues (2012) and Carlson and Wu (2012) noted that when researchers did justify their control variables, they typically did so by noting the statistical association between the predictor and the control (e.g., the predictor and third variable correlate at .40, so it is appropriate to control for the third variable).

Psychological researchers also receive relatively little helpful advice about what constitutes strong evidence for inclusion of a control variable in their area of research. The available advice is often too vague (e.g., “offer rational explanations, citations, statistical/empirical results, or some combination”; Becker, 2005, p. 278), too minimalistic (e.g., use a theoretical model to motivate control variable selection; Breugh, 2006), or is of little practical use (e.g., provide evidence that control variables are accomplishing their intended purpose; Carlson & Wu, 2012). The handful of articles that have focused more specifically on the need to explicate relations between the control, predictor, and outcome provide some helpful guidance, but it can be difficult to know how to implement this guidance in one’s own line of research. For example, Meehl (1970) noted that controls should not be automatically considered exogenous, and instead, researchers must consider the possibility that other important variables in the model—the predictor or outcome—might affect the control. In addition,

others have argued that researchers should outline the theory behind their decision to include/exclude control variables (Bernerth & Aguinis, 2016; Edwards, 2008). However, on their own, these calls to integrate theory may be difficult to implement. Fortunately, recent work on the importance of causal language and causal thinking in psychology (Dablander, 2020; Grosz et al., 2020; Rohrer, 2018) suggests that one way to implement these calls for proper control variables is to give more consideration to how control variables are *causally* linked to the other variables in the model.

As we show in the coming sections, when a central goal of a regression analysis is to learn about a process, the only way to qualify a variable as a good control is to consider the causal model that connects the control, predictor, and outcome. In the following sections, we show how, depending on the causal status of the third variable and the strength of its causal relations with both the predictor and outcome, controlling for the third variable can either remove or add substantial bias to the estimate of the causal effect. In doing so, we present a framework for principled control variable selection and justification.

### One Step Forward, Two Steps Back: Controlling for the Wrong Variable

Although it is possible to remove bias from an estimate of a causal path by controlling for a confounding variable, it is also easy to add bias to an estimate by controlling for a variable that is either not a confounder or does not block a confounding path (VanderWeele, 2019). In the following paragraphs, we describe different kinds of variables and discuss the consequences of controlling for each. Our goal is not to give a comprehensive list of all variables that one might possibly control for (for a more comprehensive list of “good” and “bad” control variables, see Cinelli et al., 2020) but, rather, to underscore how the impact of statistical control depends on the type of control variable. Figure 3 depicts eight types of third variables ( $C$ ) that differ in their causal relation to the predictor ( $X$ ) and outcome variable ( $Y$ ). Figure 3a shows a confounder, and Figures 3b and 3c show two “confound-blockers,” all of which can remove bias when controlled for.<sup>6</sup> Figures 3d through 3h show colliders, mediators, and a proxy, all of which are problematic when controlled for (Cinelli et al., 2020; Elwert & Winship, 2014; Pearl, 2009; Rohrer, 2018).

#### Confounder and confound-blocker

A confounder is a variable that is a (direct or indirect) cause of both  $X$  and  $Y$  (see Fig. 3a).<sup>7</sup> By controlling for a confounder, one can block the confounding path that obscures the causal effect of  $X$  on  $Y$ . But it is also possible to block the confounding path by controlling



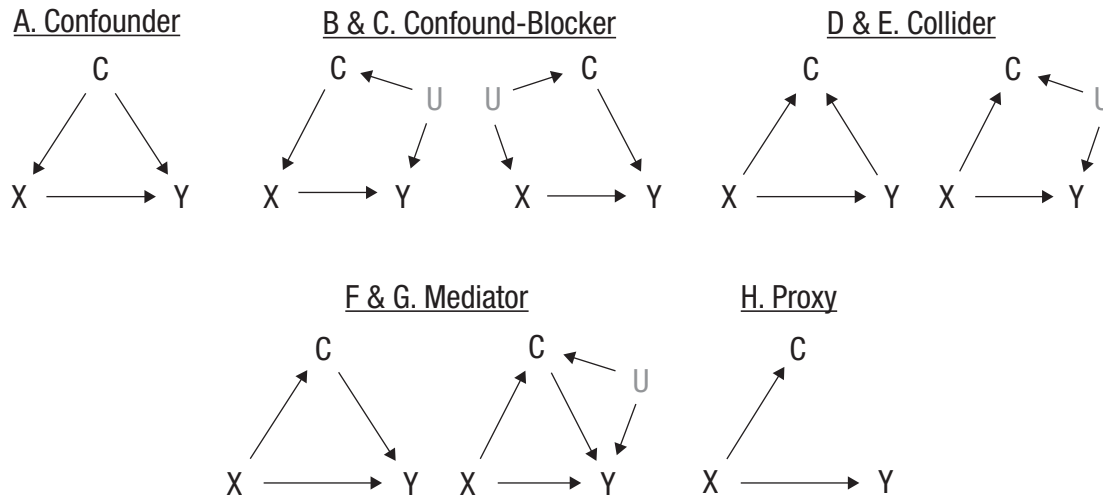


Fig. 3. Example causal models.

for any other variable that lies on that path. We call such a variable a *confound-blocker* because it is not itself a confounder, but controlling for it nevertheless blocks the confounding path. For example, to estimate the causal effect of coffee on concentration, it may be important to control for the confounding effect of sleep (because less sleep may lead to both greater caffeine consumption and lower concentration). This confounding path can be blocked either by measuring and controlling for the confounder itself—hours of sleep—or by measuring and controlling for another variable along the confounding path (e.g., desire for coffee). In Figures 3b and 3c, the confounder is unmeasured, but controlling for  $C$ , a confound-blocker, debiases the association.

### Collider

When two variables share a common effect, the common effect is called a collider between that pair of variables (Figs. 3d and 3e). For example, both IQ and hard work can result in getting accepted to college, so college-student status is a collider between IQ and hard work (Elwert & Winship, 2014; Rohrer, 2018). Controlling for a collider will induce a spurious (i.e., noncausal) association between the variables that are causes of the collider. For example, if regressing hard work on IQ produced a simple regression coefficient of zero, controlling for college-student status (which is positively affected by both hard work and IQ) would induce a spurious negative effect between these variables. A variable that is a collider for a pair of variables other than the outcome and predictor can still bias the target causal estimate. Because controlling for a collider induces a spurious association between its causes, this can transform a path between the predictor and outcome from a path that does not transmit association to a path that does. For example, in

Figure 3e,  $C$  is a collider for  $X$  and  $U$ . When  $C$  is not controlled for, the noncausal path from  $X$  to  $Y$  ( $X \rightarrow C \leftarrow U \rightarrow Y$ ) does not transmit an association because of the inverted fork ( $X \rightarrow C \leftarrow U$ ). But controlling for  $C$  induces a spurious association between  $X$  and  $U$ , and the new noncausal path from  $X$  to  $Y$  ( $X - U \rightarrow Y$ , where  $X - U$  denotes a spurious association) now transmits association, which results in a biased causal estimate.

### Mediator

A mediator is a variable that is caused by  $X$  and is a cause of  $Y$  (Figs. 3f and 3g; Baron & Kenny, 1986; Hayes, 2009; Judd & Kenny, 1981). For example, sleep problems might mediate the relation between anxiety and tiredness such that sleep problems are a mechanism by which anxiety increases tiredness. If a researcher is interested in the total effect of the predictor ( $X \rightarrow Y$  plus  $X \rightarrow C \rightarrow Y$ ) on the outcome (compared with only the direct effect,  $X \rightarrow Y$ ), then controlling for a mediator will undermine this effort by blocking one causal path of interest. Even if a researcher is interested only in the direct effect, controlling for a mediator could induce bias if the mediator and the outcome share a common cause (Fig. 3g). When such a common cause exists, the mediator is a collider for the predictor and this common cause, and because the mediator is being conditioned on, the noncausal path,  $X - U \rightarrow Y$ , now transmits association and biases the estimate (Rohrer et al., 2021).

### Proxy

A proxy is caused by  $X$  and has no causal relation to  $Y$  (Fig. 3h; Pearl, 2009). Note this does not mean the proxy is a good or sensible measure of the predictor. For example, grade point average (GPA) and number of cars

owned might be proxies for cognitive ability such that cognitive ability is a cause of GPA and (indirectly via income) cars owned. The number of cars owned, however, is likely a poor measure of cognitive ability.

If the predictor is a perfectly reliable variable (i.e.,  $X$  contains no measurement error), controlling for a proxy will not affect the  $X \rightarrow Y$  path: The regression coefficient of  $X$  will capture the causal effect, and the coefficient of the proxy regressed on the outcome will be zero (in the population). But if  $X$  is in fact an unreliable measure of the true causal variable (e.g., cognitive ability is measured with a test that is not perfectly reliable), then controlling for a proxy will attenuate the estimated causal path of interest. This attenuation effect arises because the proxy can be understood as a second unreliable measure of the same underlying predictor (e.g., GPA and the unreliable cognitive ability test are both measures of cognitive ability). When both the predictor and control variable are unreliable measures of the same construct, the true predictive effect of the construct gets partitioned into two coefficients, neither of which capture the full causal effect. The magnitude of the attenuation depends on the strength of the paths from the true (latent) predictor to the measured predictor and to the proxy.

### Inappropriate Control Leads to Bias: Demonstrating the Importance of the Causal Structure

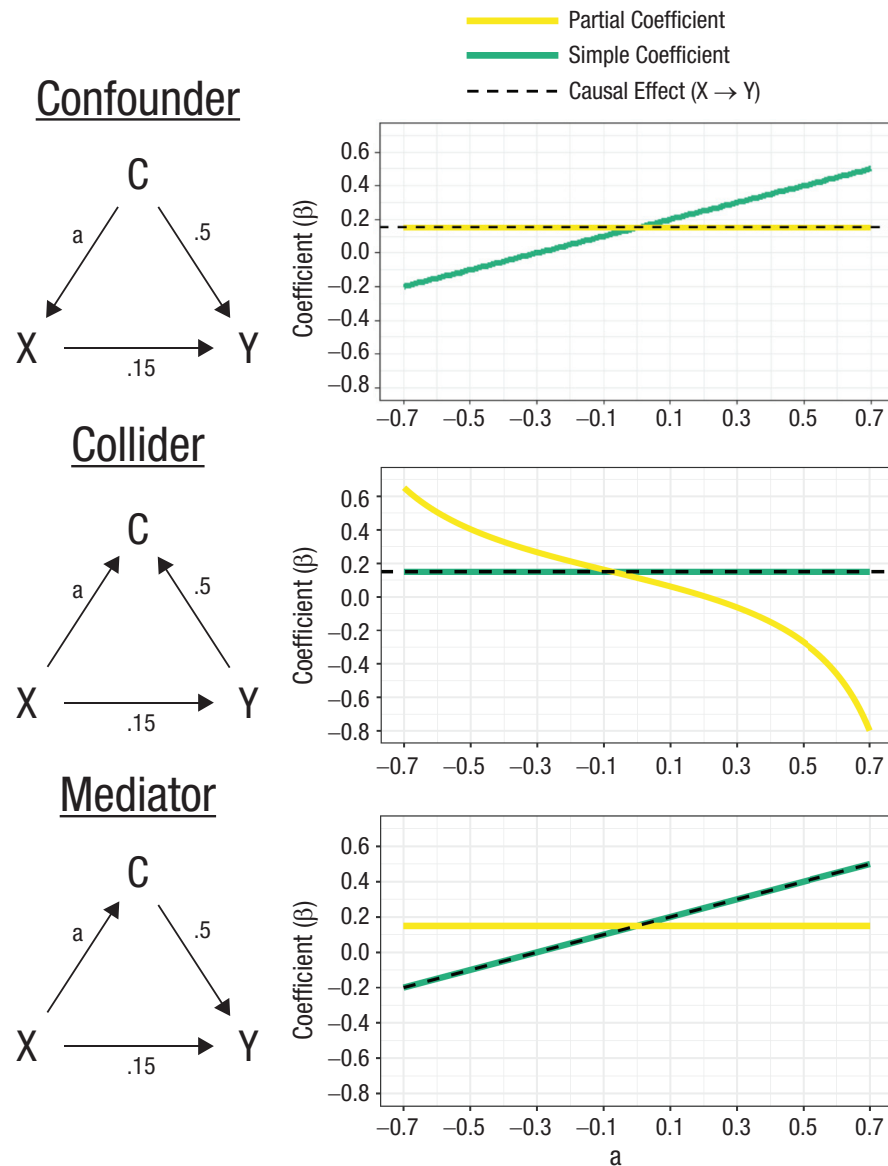
In the following section, we demonstrate how the causal structure influences the partial regression coefficients. Each figure displays the consequences of controlling for a third variable given a range of hypothetical population models. We used path tracing (i.e., Wright's rules; Alwin & Hauser, 1975; see Appendix A for an example) to obtain a population correlation matrix for each causal structure and calculated regression coefficients from each population correlation matrix using the formula  $\beta = \Sigma_{xx}^{-1} \Sigma_{xy}$ , in which  $\Sigma_{xx}$  is the  $p \times p$  correlation matrix of predictors and  $\Sigma_{xy}$  is a  $p \times 1$  vector that contains correlations between each predictor and the outcome.<sup>8</sup> We compare the partial regression coefficient (when a variable is controlled for) with the simple regression coefficient (when no other variables are controlled for). We show the population values of simple and partial regression coefficients under three of the causal structures depicted in Figure 3: when the third variable is a (a) confounder, (b) mediator, and (c) collider for the predictor and outcome (see Appendix B for results from all models shown in Fig. 3). We assume, for the time being, that all variables are measured without error (in the section, Measurement Error Makes Proxy Variables Problematic, we relax that assumption and look at the impact of controlling for a proxy). For each population model, the direct effect of  $X$  on  $Y$  is set at .15, and the direct

effect between  $C$  and  $Y$  is set at .5 (these are standardized values). The direct effect between  $X$  and  $C$  varies to demonstrate the variability of the bias within a population model.

Results are shown in Figure 4. The top row shows the effect of failing to control for a confounder: The yellow line depicts the effect of  $X$  on  $Y$  (which accurately estimates the causal effect), controlling for the confounder  $C$ , whereas the green line depicts the overestimated or underestimated value when  $C$  is not controlled for. The second and third rows of Figure 4 show two situations in which controlling for a nonconfounder introduces bias. When the control variable is a mediator, as the absolute strength of the effect of  $X$  on  $C$  increases, the total causal effect increases as well, but the partial coefficient remains the same, which results in an increasing discrepancy between the coefficient and the causal effect. When the control variable is a collider for the predictor and outcome variables, as the absolute strength of the effect of  $X$  on  $C$  increases, the discrepancy between the simple and partial coefficients increases as well. Note that the direction of the bias that arises when controlling for a collider or a mediator depends on the parameters of the model. Without knowing the true causal effect values (and we may safely assume that these are unknown!), the impact of controlling for a nonconfounder is unpredictable. Thus, if researchers are unsure whether the variable that they plan to control for is a confounder or a variable that blocks the confounding path, they should not interpret the resulting partial coefficient as a conservative approximation of the true causal effect.

### Measurement error makes proxy variables problematic

Measurement error can further muddy the interpretation of controlled regression coefficients. In the presence of measurement error, simple regression coefficients (confounded or not) will be attenuated (Shear & Zumbo, 2013), and controlling for an imperfectly measured confound may not remove its full confounding effect (Westfall & Yarkoni, 2016). Figure 5 shows results from the same series of three models as Figure 4 plus the proxy model, in which both  $X$  and  $C$  have 80% reliability. The path values that connect each of the variables are kept the same as Figure 4 except that we set the direct causal effect of  $X$  on  $Y$  in the proxy model to .5 instead of .15 to more clearly show the attenuation effect that ensues as a result of controlling for a proxy. As Figure 5 shows, controlling for a proxy is problematic when the predictor and proxy are highly correlated. As described earlier, when  $X$  is measured with error (so that the observed variable is  $X_m$  rather than  $X$  itself),  $X_m$  and  $C$  can be seen as two imperfect measures of the same true construct (with  $C$  being a much weaker measure

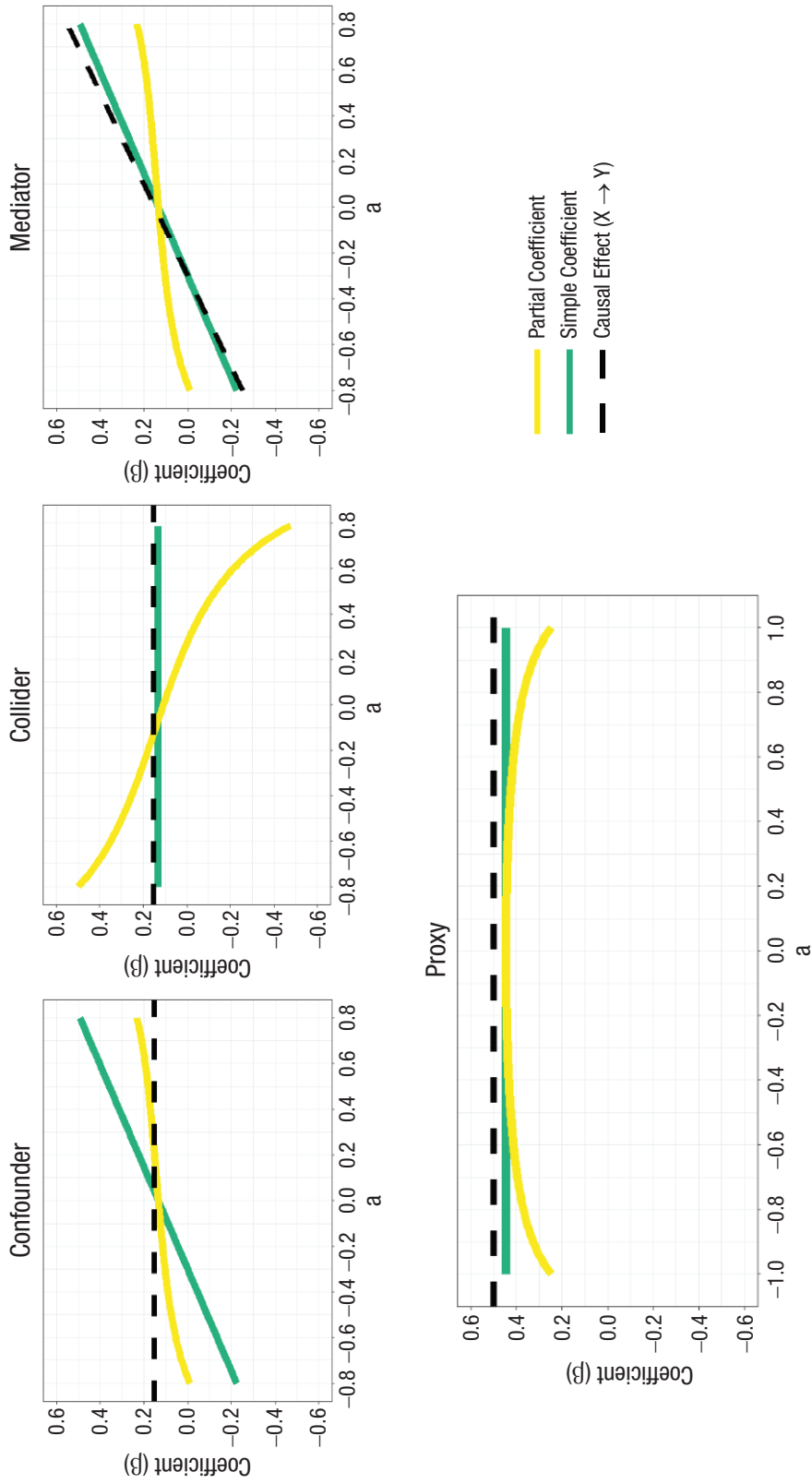


**Fig. 4.** Partial and simple regression coefficients under three causal structures. In each graph, the  $x$ -axis depicts the population value of the direct effect that connects the control variable and the predictor ( $a$ ), and the  $y$ -axis depicts the value of the regression coefficient of  $Y$  on  $X$ . The direct effect of  $X$  on  $Y$  and the value of the direct effect connecting  $Y$  and  $C$  are held constant across the results. Solid lines represent the partial (yellow line) and simple (green line) regression coefficients. The dashed line represents the total  $X \rightarrow Y$  causal effect.

than  $X_m$  to the extent that the causal path from  $X$  to  $C$  is smaller than the path from  $X$  to  $X_m$ ), and they each can account for some of the true causal effect.

Each of the causal structures in Figures 4 and 5 produces a correlation between the third variable and both the predictor and the outcome variable. In fact, the very same correlation matrix (and thus, the very same set of regression coefficients) could be produced by every one of these models. Thus, these correlations alone cannot reveal whether the third variable is, for example, a

mediator or a collider for the predictor and outcome variable (Maxwell & Cole, 2007; Pearl, 1998). Evidence of a statistical association among a third variable, predictor, and outcome merely implies that there is some causal structure that connects these variables (either directly or via a set of unobserved variables)—but the confounder structure is just one possibility among many. Therefore, it would be a mistake to assume that a variable should be controlled for merely on the grounds that it is correlated with both the predictor and outcome.



**Fig. 5.** Controlled and uncontrolled regression coefficients when variables are measured with error. In each graph, the x-axis depicts the population value of the direct effect that connects the control variable and predictor ( $a$ ), and the y-axis depicts the value of the regression coefficient of  $Y$  on  $X_m$ . Solid lines represent the partial (yellow line) and simple (green line) regression coefficients. The dashed line represents the total  $X \rightarrow Y$  causal effect. Predictor and control variables are measured with error (reliability = .8) so neither the simple nor partial coefficients capture the causal effect.



**Box 1.** Bidirectional Causality

Imagine a researcher is interested in the effect of student content knowledge on final grades and the researcher posits that the number of questions a student asks is a potential confounder for this effect. However, the researcher recognizes that the number of questions a student asks could also be a mediator. Therefore, it is plausible that (a) how much students know could influence how many questions they ask in class but also (b) how many questions students ask in class could also influence how much they know about class content. Thus, there may be bidirectional causality between the number of questions asked in class and the amount a student knows about class content. Figure 6 (left) shows one depiction of such a causal model. But this figure is misleading because it suggests that course knowledge could be a cause of questions asked and, in the same instantiation, questions asked could be a cause of course knowledge. This is impossible because causality happens across time. Thus, one of these variables—either content knowledge or questions asked—had to occur before the other, which excludes the previous variable from being caused by the latter variable. Therefore, an appropriate directed acyclic graph (DAG) should depict this bidirectional effect across different instantiations of content knowledge and questions asked (see Fig. 6, right). For example, knowledge at Week 1 of the course may be a cause of the number of questions asked in midterm week, and questions asked at Week 1 may be a cause of knowledge at midterm week. Specifying the DAG with different instantiations of the same variable allows researchers to query which causal effect they are interested in (the effect of content knowledge at Week 1 on final grade or the effect of content knowledge at midterm week on final grade) and explore which instantiation of the control variable blocks a confounding path.

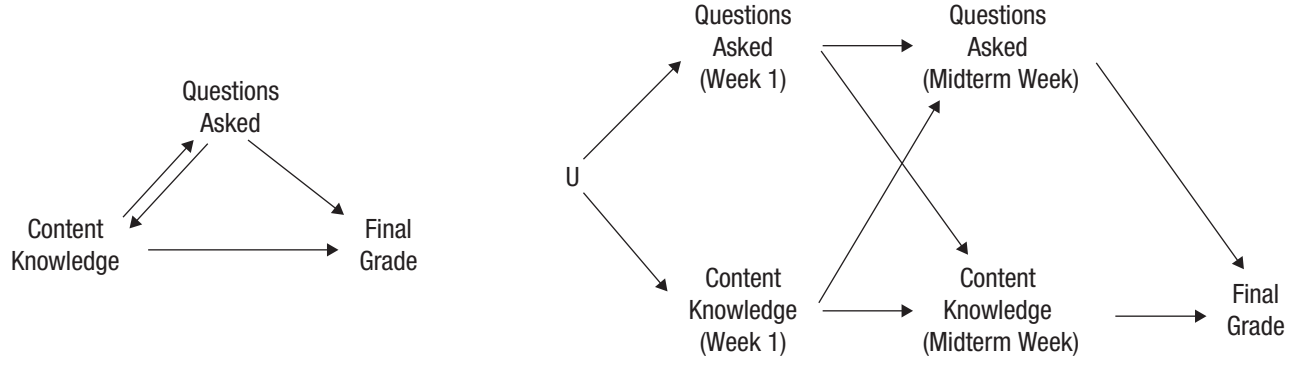
**More Complicated Models and Longitudinal Data**

In the previous section, we used simple causal structures to show how bias arises, but often, the true causal diagrams are more complicated. For instance, a causal effect may be confounded by a large set of variables, of which many are unmeasured. Measuring all confounders is not necessary if there is a more proximate variable through which many (or all) of the confounders influence the outcome or predictor. Controlling for such a variable would block all the confounding paths in which it functions as a mediator (between the confounder and the outcome or predictor) without having to measure or control for the confounders themselves.

Another complicated situation is when a potential control variable occupies two roles. For example, a variable may act as a confounder between two other constructs if measured at one time point and a mediator if measured at another; however, it is not necessarily possible to make this distinction for the same instantiation of the predictor and outcome (see Box 1). Because the goal of statistical control is to remove a confounding effect without blocking the causal effect, it is important for a researcher to identify and measure the control variable at a time when it serves as a confounder between the predictor and outcome rather than a time when it serves as a mediator. Box 1 explains in more detail what we mean and gives an example of how to do this. Likewise, if there were bidirectional causality between a control variable and the outcome, the control

variable would be a confounder if measured at some time points and a collider for the predictor and outcome variables at others. When such complicated structures exist, it can be difficult to obtain a set of control variables that debiases a causal effect. Often, longitudinal data can help with this endeavor.

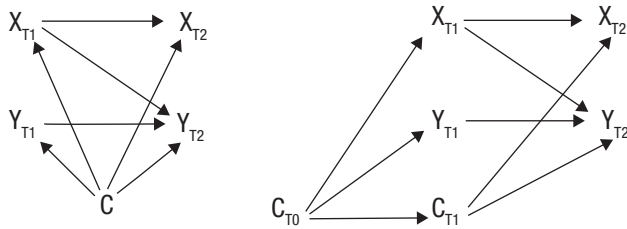
Longitudinal data—when the same variables are measured at multiple measurement occasions in the same individuals—provide information about the temporality of variables. These data, along with the use of longitudinal models, can be used to make propositions about the location and direction of effects. For example, if  $X_{T1}$  (the subscript denotes the measurement occasion) is found to predict  $Y_{T2}$  even after controlling for previous measurements of  $Y$ , then  $X$  is said to Granger-cause  $Y$ . Granger causality depends on two criteria: (a) The Granger-cause precedes its effect, and (b) the Granger-cause explains unique variation in its effect over and above what is predicted by a previous measure of  $Y$  (Granger, 1980; Maziarz, 2015). Establishing Granger causality is not the same thing as establishing causality, however (Eichler & Didelez, 2010). Effects that meet the definition of Granger causality may still be influenced by confounders because controlling for a previous version of  $Y$  may not block all confounding paths. For example, in Figure 7 (left), controlling for  $Y_{T1}$  blocks one of the confounding paths ( $X_{T1} \leftarrow C \rightarrow Y_{T1} \rightarrow Y_{T2}$ ) but not the other ( $X_{T1} \leftarrow C \rightarrow Y_{T1}$ ). Longitudinal data can make it much easier to control for confounders, but it does not negate the need to clearly justify the underlying causal structure (Rohrer, 2019).



**Fig. 6.** Incorporating multiple instantiations of a variable into a directed acyclic graph.  $U$  represents a set of unspecified variables that causes both the number of questions asked and amount of content knowledge in Week 1.

Combined with a justified causal structure, longitudinal data can address some complicated causal structure problems. First, by repeatedly measuring a pair of variables that influences each other ( $X \rightarrow C$  and  $C \rightarrow X$ ) at the correct interval, the once bidirectional paths become unidirectional paths (e.g.,  $X_{T1}$  is a cause of  $C_{T2}$ , and  $C_{T1}$  is a cause of  $X_{T2}$ ).<sup>9</sup> Second, measuring the same variables across time allows for the removal of a specific kind of unmeasured confounding, namely, *time-invariant* confounding. A time-invariant confounder (Fig. 7, left) is a confounder whose level and effects do not change across the measured time points (e.g., ethnicity—assuming the effect of participants’ ethnicity on the other variables in the model is constant across measurement occasions). In contrast, a time-varying confounder (Fig. 7, right) is a variable whose level or effect changes between measured time points (e.g., positive affect, relationship satisfaction).

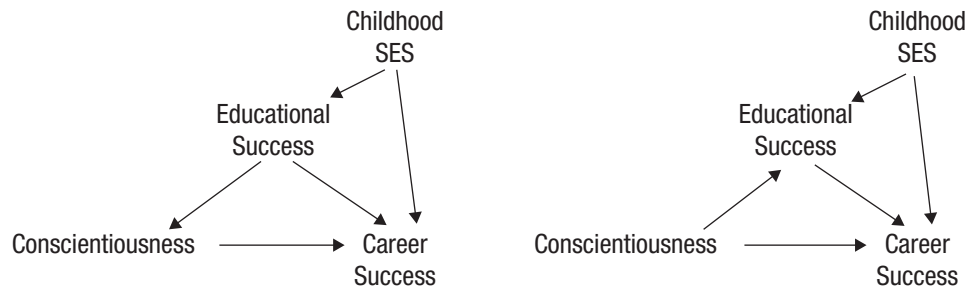
One way to remove unmeasured time-invariant confounds is to use a fixed-effects model (Allison, 2005; Kim & Steiner, 2021).<sup>10</sup> A fixed-effects model estimates



**Fig. 7.** Examples of time-invariant and time-varying confounders. (Left) The confounder,  $C$ , does not change across the two measurements. There are two confounding paths between  $X_{T1}$  and  $Y_{T2}$ :  $X_{T1} \leftarrow C \rightarrow Y_{T2}$  and  $X_{T1} \leftarrow C \rightarrow Y_{T1} \rightarrow Y_{T2}$ . Controlling for  $Y_{T1}$  blocks only the second confounding path. (Right) The confounder does change across measurements. Because an effect precedes its cause, the subscript for the first version of the confounder is 0. Again, there are two confounding paths:  $X_{T1} \leftarrow C_{T0} \rightarrow C_{T1} \rightarrow Y_{T2}$  and  $X_{T1} \leftarrow C_{T0} \rightarrow Y_{T1} \rightarrow Y_{T2}$ . Controlling for  $Y_{T1}$  blocks only the second confounding path.

an individual-specific intercept, which captures all effects that vary between but not within individuals in a sample (or another unit of analysis such as school or country). Because time-invariant confounds do not vary within individuals, the fixed-effects model can debias a coefficient for unmeasured time-invariant confounders. A fixed effect can be estimated by including a dummy variable in the regression model for each participant (for tutorials on estimating a fixed-effects model in R, see Colonescu, 2016; Hanck et al., 2019). Rather than controlling for a specific variable, the fixed-effects model simultaneously controls for all the attributes of individuals that do not vary over time. Although fixed-effects models remove the effects of unmeasured time-invariant confounders, they do not remove the effects of *time-varying* confounders. Therefore, considering whether time-varying confounders exist for a pair of variables and making a plan, if needed, to deal with them is still important when a fixed-effects model is estimated.

In summary, the causal structure must still be proposed and justified even if longitudinal data are available (Hernán et al., 2002; Imai & Kim, 2016), and the only way to provide a valid argument that it is appropriate to control for a particular variable is to discuss the causal structure that includes the control, predictor, and outcome. This argument, which should be presented for each control variable, can include empirical motivation (e.g., an estimated association from a previous study), but it must be justified on a theoretical basis (outlined below). This strategy aligns with advice from psychometricians to be conservative with the number of control variables (Becker et al., 2016; Carlson & Wu, 2012) and to carefully justify each one (Bernierth & Aguinis, 2016; Breauh, 2006; Carlson & Wu, 2012; Edwards, 2008). Of course, proposing a full causal map of relations among all the variables in one’s model is more difficult than using the default methods of



**Fig. 8.** Considering alternative structures: educational attainment as a confounder or a mediator.

selecting control variables. To assist in this difficult but necessary endeavor, we next use an applied-research example to demonstrate how to decide which variables to control for using causal reasoning.

### Using the Causal Structure to Justify Control Variables: An Applied Example

In this section, we outline the steps to properly justify control variables using causal structures. As a demonstration, we use a simplified, applied example from the literature on personality and work.

#### Selecting variables

Begin with a pair of outcome and predictor variables, in which the predictor variable is hypothesized to be a cause of the outcome.<sup>11</sup> We chose conscientiousness, or the disposition to be hardworking, responsible, and organized, as the predictor and career success, including annual income, occupational prestige, and job satisfaction, as the outcome. Past research supports conscientiousness as a predictor of future career success (Dudley et al., 2006; Moffitt et al., 2011; Sutin et al., 2009; Wilmot & Ones, 2019). We propose that this relation exists because being a conscientious worker (e.g., fulfilling work responsibilities, being punctual, working diligently) increases one's career success.

Next, generate a list of variables that could be confounders or confound-blockers—the (potential) controls. This list should include variables that are statistically associated with both predictor and outcome. We considered variables that might be confounders of personality and career success and identified two—educational attainment and childhood socioeconomic status (SES).

#### Specifying and justifying the causal structure

With this list of potential controls, the next step is to specify a causal structure that includes the predictor,

outcome, potential confounders, and other important third variables (for our example, see Fig. 8). For this step, researchers can use *U* to stand in for a set of non-specific common causes. Using *U* as a stand-in variable allows researchers to reason about potential confounding paths and consider how to block them even when the full set of confounders is unknown. After outlining a causal structure, each part of the structure should be justified. For the list of potential controls, researchers must justify why these variables are either confounders or confound-blockers of the predictor and outcome. For example, according to the social investment principle of personality, which posits that age-graded social roles serve as one mechanism of personality development (Roberts & Wood, 2006), we hypothesized that educational attainment influences conscientiousness, that is, engaging in the structured context of higher education that demands individuals to act responsibly should promote individuals to become more organized, hardworking, and responsible (i.e., more conscientious; Roberts et al., 2004). We also theorized that higher educational attainment would lead to greater career success given robust associations between educational attainment and income, unemployment, job satisfaction, occupational prestige, and control over work (Gürbüz, 2007; Ross & Reskin, 1992; Slominski et al., 2011; U.S. Bureau of Labor Statistics, 2020). Together, these two lines of research support our hypothesized structure that educational attainment is a plausible confounder of the causal effect of conscientiousness on career success.

In addition, childhood SES may also be a confounder or confound-blocker for conscientiousness (mediated by educational attainment) and career success because children from families with higher incomes and more highly educated parents are, themselves, more likely to achieve higher levels of educational attainment and have higher paying and more prestigious jobs (National Center for Education Statistics, 2019; The Pell Institute, 2018). These causal hypotheses are depicted in Figure 8 (left). Of course, any of these hypotheses about the causal structure may be wrong. The value of this framework, however, lies in the clear outlining of assumptions and

hypotheses and, in turn, the consequences that arise if these assumptions are wrong.

Researchers should also justify why each control variable will not bias the estimate by blocking a causal path (i.e., it is not a mediator) or inducing a spurious path. In most research studies, there will be uncertainty about parts of the causal structure (e.g., a variable may be a plausible confounder and also a plausible collider for a pair of variables). These uncertainties can be depicted by having multiple competing models—a set of plausible causal structures. In our example, there is reason to believe that educational attainment might be a mediator between conscientiousness and career success, although it is unlikely to be a proxy or a collider.<sup>12</sup> That is, conscientiousness may be a cause of educational attainment (because being dispositionally hardworking and reliable causes people to do better in school and receive more opportunities to further their education; Göllner et al., 2017), and higher educational attainment, in turn, causes career success. Childhood SES, however, could be neither a mediator nor a collider for conscientiousness and career success because events that occur in adulthood do not change events that happened in childhood. Therefore, we propose that we have a clear confounder—childhood SES—and one variable that could be a confounder and/or a mediator—educational attainment.

### ***Comparing alternative models and selecting control variables***

With a set of plausible causal structures, the next step is to select an appropriate set of control variables, which should block all confounding paths without blocking any causal paths or inducing any spurious associations between the predictor and outcome. This appropriate set of control variables need not contain every confounder—sometimes all confounding paths can be blocked with a subset of the confounders.

With a set of appropriate controls for each plausible model, researchers will end up in one of three positions. First, there may be a set of control variables that is appropriate across all plausible models. In this case, the researcher can argue that the value of the partial coefficient is a more accurate estimate of the hypothesized causal path than the simple coefficient. Researchers who interpret their estimated association causally must make it clear to readers that the causal interpretation of the results is predicated on the specified model being true and the assumptions outlined previously (see Statistical Control: How and When It Works section; e.g., all non-linear and interaction effects are correctly specified). Second, researchers may find that the appropriate set of control variables varies across the set of plausible models. In this case, researchers could select a single model and control for its corresponding control set and discuss

how the interpretation of the results depends on the chosen structure being true. Here, the alternative models should be included in the article along with their associated control sets. Alternatively, researchers could run separate models that control for each of the appropriate control sets and present the partial associations from each model along with a discussion of the assumptions that have to hold for each of these coefficients to be unbiased for the causal effect. Researchers may also consider conducting a sensitivity analysis to investigate which effects hold up to control by various possible confounders, keeping in mind that without making assumptions about the causal model, there is no basis to claim that partial effects are more or less “conservative” than simple effects. Finally, researchers could be in a position in which one or more of the models have no appropriate control set. When this is the case, researchers may consider blocking the confounding paths through some other method (e.g., instrumental variables or front-door criterion; Pearl, 1995), reporting the simple coefficients, or reporting the partial coefficients with an acknowledgment that not all of the confounding paths are blocked. These recommendations are summarized in Table 1.

When reporting partial coefficients, there are some practices researchers should always follow. First, both the simple and partial coefficients should be made accessible to the reader; not providing access to both sets of coefficients leaves the reader without valuable information about how statistical control affects the estimates. In addition, it is important to stress that the coefficient that relates the outcome to the predictor and the coefficients that relate the outcome to the controls cannot generally be interpreted in the same way because control  $\rightarrow$  outcome coefficients represent direct (rather than total) effects and may themselves be confounded<sup>13</sup> (Westreich & Greenland, 2013).

In our example, because educational attainment may serve as either a mediator or a confounder, we must consider and discuss the consequences of both causal models (see Fig. 8). If educational attainment is a confounder, then controlling for educational attainment would block the two confounding paths. Instead, if educational attainment is a mediator, then controlling for educational attainment not only blocks a causal path but also induces a spurious path (because of Conscientiousness  $\rightarrow$  Educational Attainment  $\leftarrow$  Childhood SES). Thus, if educational attainment is a mediator, controlling for childhood SES is the correct approach. If longitudinal data were available, then it could be possible to distinguish educational attainment as a confound from educational attainment as a mediator. Longitudinal data would also allow fixed effects to be estimated, which would remove time-invariant confounding by controlling for all fixed person-level attributes. Another way to disentangle whether a variable is

**Table 1.** Guidelines for Controlling After Justification Process

Result of causal structure reasoning	Suggested approach
A single control set is appropriate across the plausible models	<ul style="list-style-type: none"><li>• Control for the control set</li><li>• List assumptions that the causal interpretation is predicated on</li></ul>
Different control sets are appropriate across plausible models	<ul style="list-style-type: none"><li>• Control for one of the control sets</li><li>• Discuss how the interpretation of the coefficient depends on the selected model being true</li><li>Or</li><li>• Run multiple models with different control sets</li><li>• Present results from different models as competing estimates</li></ul>
No control set	<ul style="list-style-type: none"><li>• Use another method to debias effects (e.g., instrumental variables)</li><li>Or</li><li>• Report simple coefficients</li><li>Or</li><li>• Report partial coefficients with an acknowledgment that it is a biased coefficient</li></ul>

a confounder or a mediator is to take a more nuanced view of each of the variables (see Box 2).

Discussion

In this article, we demonstrated the importance of carefully selecting control variables. In particular, we highlighted how controlling for the wrong variable can lead researchers to results and interpretations that are less accurate than if no variables had been controlled. Furthermore, we showed that the underlying causal structure determines whether controlling for a variable adds or removes bias. In addition, we clarified that statistical associations are not sufficient justification for selecting a control variable because these associations could arise from a number of different causal structures.

Throughout this article, we discussed estimating the weights of causal paths and how these weights can be biased by controlling for the wrong variable (researchers

may also be interested in establishing incremental validity; see Box 3 for a discussion of this practice). But this framework is important even for researchers who are interested only in assessing whether a causal effect exists (rather than estimating the weight of that causal effect). In this case, the researcher may not be concerned whether the weight of the causal path is an underestimate (or overestimate) as long as the results indicate that there is a non-zero causal path between the two variables. There are two reasons why the causal structure is important even if the existence, rather than the weight, of a causal effect is the focus. First, controlling for the wrong variable can, in some situations, entirely remove the effect of interest. In particular, if the impact of *X* on *Y* is mainly mediated by a third variable and that mediator is controlled for, the association between *X* and *Y* can be reduced to (or very near to) zero, which leads to the relation being mistakenly dismissed as unimportant. The same thing can happen if a spurious association that biases the effect of interest is induced

Box 2. Breaking Down Variables

Figure 8 depicts directed acyclic graphs (DAGs) that include variables that are amalgams of complex processes that unfold across time (e.g., educational attainment). Collapsing over a variable can be useful, but it can also create ambiguity when debiasing a causal effect. One way to disambiguate this process is by specifying each variable’s content and timescale (see Fig. 9). For example, educational attainment could be broken down into multiple parts—being a college attendee, engaging in the educational demands of college (e.g., completing schoolwork), and receiving a college degree. In addition, conscientiousness can be assessed multiple times—once during high school and once after high school. Adding this nuance helps to clarify the causal effects; in particular, early conscientiousness influences college admittance and the educational demands of college influences later conscientiousness. It also clarifies how childhood socioeconomic status affects multiple steps in the educational path, including college attendance and degree attainment. Assuming this more nuanced structure is accurate, controlling for childhood socioeconomic status would allow for the effect of post-high-school conscientiousness on later career success to be identified.



### Box 3. What About Controlling to Assess Incremental Validity?

Researchers will sometimes cite incremental validity, rather than explanation, as the motivation for statistical control. Incremental validity means that a new predictor accounts for substantial variance in an outcome variable over and above the variance accounted for by other, known predictors (Sechrest, 1963), and it is typically established using multiple regression with the previously known predictors included as controls. If the new predictor accounts for significant variance over and above the other predictors, then the new predictor is said to have incremental validity.

There are three common ways in which incremental validity is used in psychological research. The first is to establish the clinical usefulness of a test to predict a specific criterion (Smith et al., 2003). In this use, the research goal is improved prediction, and it is appropriate to choose control variables that represent whatever measures are currently most popular or predictive in the field. Incremental validity for the sake of establishing predictive utility does not require causal justification.

The second common use of incremental validity is to establish that a new theoretically motivated predictor explains a phenomenon over and above previously established predictors. In this case, incremental validity allows researchers to “[demonstrate] the empirical novelty of a conceptual contribution” while “simultaneously [acknowledging] a precedent established in earlier work” (Wang & Eastwick, 2020, p. 157). This use of incremental validity can lead to greater “understanding of the relations between a focal predictor, a covariate, and an outcome variable” (Wang & Eastwick, 2020, p. 157). We have endeavored to make clear by now that any method that aims to build up a theoretical understanding of variable relations needs to consider the causal direction of those relations when choosing which predictors to control for.

But Wang and Eastwick (2020) described a third use of incremental validity testing, which may initially seem not to require causal theory, namely, the “isn’t-it-just” argument. The isn’t-it-just argument uses incremental validity to dispute a criticism that a new measure predicts only an outcome because it is a proxy for some already established predictor (e.g., “You found an effect of Machiavellianism on aggressive behavior, but isn’t Machiavellianism just psychopathy?”). This “isn’t-it-just” argument, although it uses different language, is equivalent to ruling out an alternative explanation that Machiavellianism is a proxy for psychopathy (i.e., variability in psychopathy causes or produces variability in scores on a Machiavellianism scale). Someone who argues this point might claim that the estimated effect is due to a causal path from psychopathy to aggressive behavior rather than such a path from Machiavellianism to aggressive behavior (see Fig. 10). In other words, the “isn’t-it-just” argument is another way to posit a confounder. Thus, it makes sense to control for psychopathy to estimate the remaining (i.e., unique) effect of Machiavellianism on aggressive behavior. But before testing for incremental validity, it is still important to establish that psychopathy is a potential confounder by defending the underlying causal structure. If it were more plausible that psychopathy was a proxy for Machiavellianism, rather than the other way around, or that psychopathy was a plausible mediator of the effect of Machiavellianism on aggressive behavior, then it would not make sense to investigate the incremental validity of Machiavellianism over and above psychopathy. Therefore, researchers making the “isn’t-it-just” argument to test for incremental validity should carefully consider the plausible causal ordering of variables to appropriately control for third variables.

when controlling for a collider, and this association is of the opposite sign and sufficiently strong. In addition, although researchers may not be concerned with the exact

weight of the path of interest, they likely care that it is at least in the right direction (e.g., if higher Machiavellianism increases aggressive behavior, then it would not be helpful

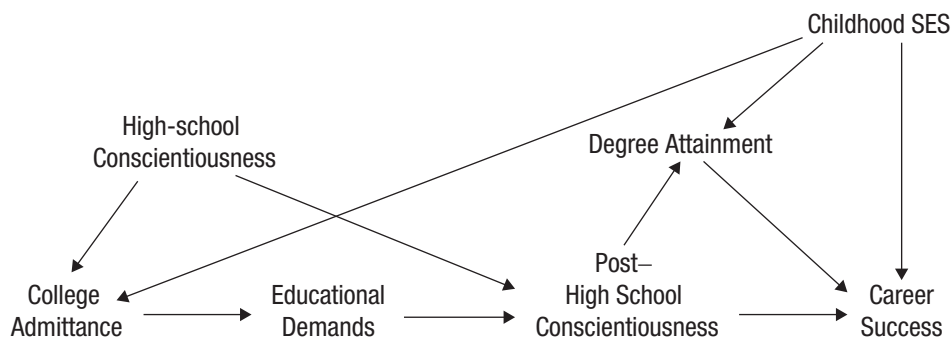
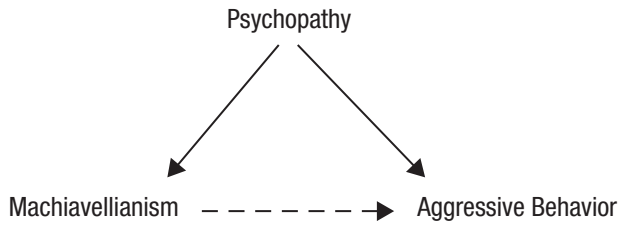


Fig. 9. Breaking down variables in a directed acyclic graph.



**Fig. 10.** Incremental validity as a question of confounding.

to have results indicating that higher Machiavellianism decreases aggressive behavior). As Figures 4 and 5 show, controlling for an inappropriate third variable can inaccurately flip the sign of the coefficient, which leads to an estimate that indicates a causal effect in the opposite direction. Therefore, we reiterate that for controlled results to be meaningfully interpreted to explain a process, a causal structure must be proposed and defended. Without a causal structure, neither the researcher nor the reader can make sense of discrepancies between the partial and simple coefficients.

## Conclusion

Although the causal structure holds the key to identifying confounders, it is uncommon for psychological researchers to present causal justification for their choice of control variables. The absence of explicit causal reasoning may be due to an unspoken ban on causal language within psychology (Dablander, 2020; Grosz et al., 2020). In contrast to fields such as epidemiology and economics, psychologists who report on nonexperimental findings often claim to be interested in prediction or association even though the interpretations made within an article are more compatible with causal inference (Grosz et al., 2020). In this article, we argued that statistical control for the goal of causal inference is incompatible with causal agnosticism about how the control variable relates to the predictor and outcome: Whether controlling for some variable increases or decreases bias is a function of how the variables in a model causally relate to each other.

There are three main benefits to justifying the causal structure. First, control variables are selected in a more appropriate and careful manner. This will improve results from analyses with control variables. Second, there will be an improvement in the theoretical models that motivate the study. Most researchers have, at least implicitly, a hypothesized causal structure. Taking the time to (literally) draw the causal model may give researchers an opportunity to think carefully about their causal assumptions and to consider alternative plausible causal structures. Third, when the hypothesized causal structure is made explicit, readers will more easily glean the causal framework the authors are working from. Readers can then be aware of the assumptions that are inherent in a model and can reason about how those assumptions may

influence the results if they are violated. In summary, psychological research stands to benefit substantially from researchers thinking and communicating carefully about why they selected specific control variables.

## Appendix A

Wright (1934) outlined an approach (i.e., Wright's rules) to obtain an implied correlation matrix from a causal structure or a path diagram. Assuming that all variables have variance = 1, the correlation between two variables can be found by summing all compound paths that travel between the two variables. A compound path (a) must not go through the same variable twice, (b) cannot go forward ( $\rightarrow$ ) and then backward ( $\leftarrow$ ; e.g.,  $X \rightarrow Z \leftarrow Y$  is not a valid compound path), and (c) can contain at most one double-headed arrow.

Figure A1 depicts a causal diagram for four variables ( $W$ ,  $X$ ,  $Y$ , and  $Z$ ) in which the lowercase letters "a," "b," "c," and "d" represent the causal path coefficients. The implied correlations for each pair of variables in this model are as follows:

$$\text{Cor}(W, X) = a$$

$$\text{Cor}(W, Z) = ac + abd$$

$$\text{Cor}(W, Y) = ab$$

$$\text{Cor}(X, Y) = b$$

$$\text{Cor}(X, Z) = c + bd$$

$$\text{Cor}(Y, Z) = d + bc$$

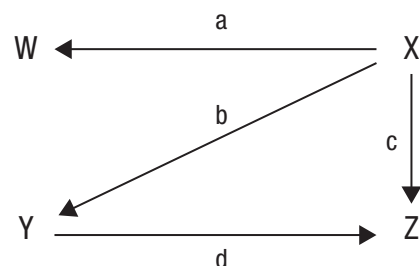
If  $a = b = c = d = .5$ , then

$$\text{Cor}(W, Z) = .25 + .125 = .375$$

$$\text{Cor}(W, X) = .5$$

$$\text{Cor}(W, Y) = .25$$

$$\text{Cor}(X, Y) = .5$$



**Fig. A1.** Path diagram depicting causal paths among variables  $W$ ,  $X$ ,  $Y$ , and  $Z$ .

$$\text{Cor}(X, Z) = .5 + .25 = .75$$

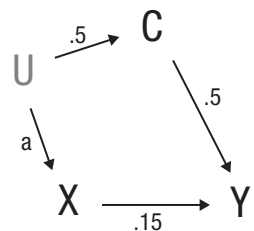
$$\text{Cor}(Y, Z) = .5 + .25 = .75$$

This correlation matrix can then be used to calculate regression coefficients (for R code, see OSF page, [https://osf.io/64rfv/?view\\_only=f49974350af14a7185994eeb00374306](https://osf.io/64rfv/?view_only=f49974350af14a7185994eeb00374306)).

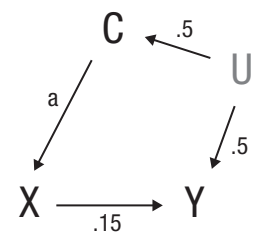
## Appendix B

(see Fig. B1)

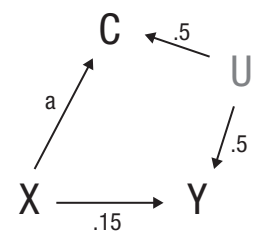
### Confound-Block



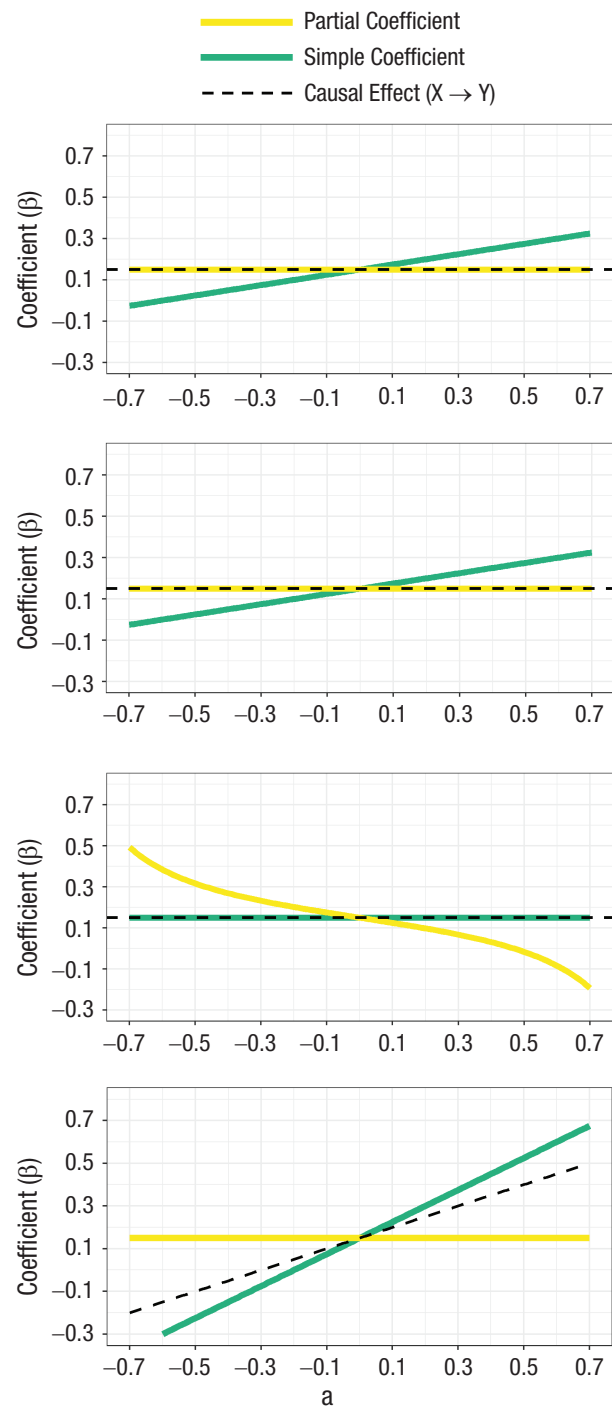
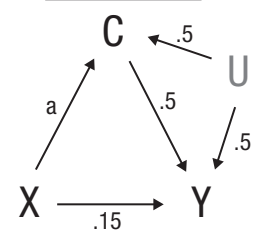
### Confound-Block



### Collider



### Mediator



**Fig. B1.** Partial (adjusted for  $C$ ) and simple (unadjusted) regression coefficients under four causal structures.

## Transparency

Action Editor: Robert L. Goldstone

Editor: Daniel J. Simons

### Author Contributions

Conceptualization: A. Wysocki, K. M. Lawson, M. Rhemtulla; supervision: M. Rhemtulla; visualization: A. Wysocki; writing, original draft: A. Wysocki, K. M. Lawson, M. Rhemtulla; writing, review and editing: A. Wysocki, K. M. Lawson, M. Rhemtulla. All of the authors approved the final manuscript for submission.

### Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

### Open Practices

Open Data: not applicable


Open Materials: <https://osf.io/64rfv/>

Preregistration: not applicable

All materials have been made publicly available via OSF and can be accessed at <https://osf.io/64rfv/>. This article has received the badge for Open Materials. More information about the Open Practices badges can be found at <http://www.psychologicalscience.org/publications/badges>.



## ORCID iDs

Anna C. Wysocki  <https://orcid.org/0000-0001-8102-2313>

Katherine M. Lawson  <https://orcid.org/0000-0002-4083-9797>

Mijke Rhemtulla  <https://orcid.org/0000-0003-2572-2424>

## Notes

1. Note that many statisticians prefer the term “adjust” over “control” in the context of regression. The concern is that statistical control may be mistakenly conflated with (the stronger) experimental control (e.g., Gelman, 2019). Although we agree that “adjust” is more precise, we continue to use the term “control” because it is more prevalent in applied psychological research.
2. We use the terms “accuracy” and “bias” to describe the relation of a population regression coefficient ( $Y$  regressed on  $X$ ) to the average causal effect of  $X$  on  $Y$ . This use is different from how “bias” is used in statistics, in which it describes the relation of an estimated statistic to its population quantity.
3. DAG Figures 1, 3, 4, 5, and 10 were created in R using the package *ggDag* (Barrett, 2021). All result figures were created in R (Version 3.6.2; R Core Team, 2021) using the package *ggplot2* (Wickham, 2016).
4. Assuming two (or more) paths do not cancel each other out.
5. Some research questions can be answered only when the set of variables is measured at a specific interval. For example, if the research question pertains to how asking questions in class affects one’s final grade, then the “question” variable should represent the number of questions each student asked across the entire class rather than just a single week.
6. Bias amplification can occur when, after controlling for a confounder or an instrumental variable (especially one that is strongly correlated with  $X$  and only weakly correlated with  $Y$ ), the bias

from a second uncontrolled confounding path can increase. If the bias removed by controlling for one confounder is less than the amount of bias induced by bias amplification, the overall bias can increase (Steiner & Kim, 2016).

7. Some other definitions of a confounder (e.g., VanderWeele & Shpitser, 2013) do not draw a distinction between confounders and confound-blockers.

8. The code used to calculate the coefficients is available at [https://osf.io/64rfv/?view\\_only=f49974350af14a7185994eeb00374306](https://osf.io/64rfv/?view_only=f49974350af14a7185994eeb00374306).

9. The interval at which a set of variables should be measured depends on the causal effect of interest (Lundberg et al., 2021). For example, the effect of high school mentorship on postcollege income is a different causal effect than the effect of college mentorship on postcollege income. Each of these causal effects would require measuring mentorship at a different interval.

10. Two other methods of removing time-invariant confounding is by estimating mean deviation scores and gain scores. See Kim and Steiner (2021) for a comparison of these methods.

11. If a predictor is not a likely cause of the outcome, then the researcher may be interested in prediction rather than explanation. In that case, they may want to choose a set of predictors that maximizes the variance explained rather than focusing on the interpretation of coefficients.

12. Educational attainment is unlikely to be a proxy for conscientiousness, given its clear and well-substantiated relation to career success. In addition, it is unlikely to be a collider between conscientiousness and work success, given the typical temporal relation between education and career success (i.e., most individuals complete their education before entering in the workforce).

13. The tendency to interpret these coefficients in the same way is known as the Table 2 Fallacy.

## References

- Allison, P. D. (2005, August). *Causal inference with panel data* [Paper presentation]. Proceedings of the Annual Meeting of the American Sociological Association.
- Alwin, D. F., & Hauser, R. M. (1975). The decomposition of effects in path analysis. *American Sociological Review*, 40(1), 37–47.
- Atinc, G., Simmering, M. J., & Kroll, M. J. (2012). Control variable use and reporting in macro and micro management research. *Organizational Research Methods*, 15(1), 57–74.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173–1182.
- Barrett, M. (2021). *ggdag: Analyze and create elegant directed acyclic graphs*. <https://CRAN.R-project.org/package=ggdag>
- Becker, T. E. (2005). Potential problems in the statistical control of variables in organizational research: A qualitative analysis with recommendations. *Organizational Research Methods*, 8(3), 274–289.
- Becker, T. E., Atinc, G., Breaugh, J. A., Carlson, K. D., Edwards, J. R., & Spector, P. E. (2016). Statistical control in correlational studies: 10 essential recommendations for organizational researchers. *Journal of Organizational Behavior*, 37(2), 157–167.

- Bernerth, J. B., & Aguinis, H. (2016). A critical review and best-practice recommendations for control variable usage. *Personnel Psychology*, 69(1), 229–283.
- Borsboom, D., van der Maas, H. L., Dalege, J., Kievit, R. A., & Haig, B. D. (2021). Theory construction methodology: A practical framework for building theories in psychology. *Perspectives on Psychological Science*, 16(4), 756–766. <https://doi.org/10.1177/1745691620969647>
- Breaugh, J. A. (2006). Rethinking the control of nuisance variables in theory testing. *Journal of Business and Psychology*, 20(3), 429–443.
- Breaugh, J. A. (2008). Important considerations in using statistical procedures to control for nuisance variables in non-experimental studies. *Human Resource Management Review*, 18(4), 282–293.
- Carlson, K. D., & Wu, J. (2012). The illusion of statistical control: Control variable practice in management research. *Organizational Research Methods*, 15(3), 413–435.
- Cinelli, C., Forney, A., & Pearl, J. (2020). *A crash course in good and bad controls*. [https://ftp.cs.ucla.edu/pub/stat\\_ser/r493.pdf](https://ftp.cs.ucla.edu/pub/stat_ser/r493.pdf)
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Erlbaum.
- Colonescu, C. (2016). *Principles of econometrics with R*.
- Cui, X., Guo, W., Lin, L., & Zhu, L. (2009). Covariate-adjusted nonlinear regression. *The Annals of Statistics*, 37(4), 1839–1870.
- Dablander, F. (2020). *An introduction to causal inference*. PsyArXiv. <https://doi.org/10.31234/osf.io/b3fkx>
- Dudley, N. M., Orvis, K. A., Lebiecki, J. E., & Cortina, J. M. (2006). A meta-analytic investigation of conscientiousness in the prediction of job performance: Examining the inter-correlations and the incremental validity of narrow traits. *Journal of Applied Psychology*, 91(1), 40–57.
- Edwards, J. R. (2008). To prosper, organizational psychology should . . . overcome methodological barriers to progress. *Journal of Organizational Behavior*, 29(4), 469–491.
- Eichler, M., & Didelez, V. (2010). On Granger causality and the effect of interventions in time series. *Lifetime Data Analysis*, 16(1), 3–32.
- Elwert, F., & Winship, C. (2014). Endogenous selection bias: The problem of conditioning on a collider variable. *Annual Review of Sociology*, 40, 31–53.
- Gelman, A. (2019, January 25). When doing regression (or matching, or weighting, or whatever), don't say "control for," say "adjust for." *Statistical Modeling, Causal Inference, and Social Science*. <https://statmodeling.stat.columbia.edu/2019/01/25/regression-matching-weighting-whatever-dont-say-control-say-adjust/>
- Göllner, R., Damian, R. I., Rose, N., Spengler, M., Trautwein, U., Nagengast, B., & Roberts, B. W. (2017). Is doing your homework associated with becoming more conscientious? *Journal of Research in Personality*, 71, 1–12.
- Granger, C. W. (1980). Testing for causality: A personal viewpoint. *Journal of Economic Dynamics and Control*, 2, 329–352.
- Greenland, S. (1990). Randomization, statistics, and causal inference. *Epidemiology*, 1(6), 421–429.
- Grosz, M. P., Rohrer, J. M., & Thoemmes, F. (2020). *The taboo against explicit causal inference in nonexperimental psychology*. PsyArXiv. <https://doi.org/10.31234/osf.io/8hr7n>
- Gürbüz, A. (2007). An assessment of the effect of educational level on the job satisfaction from the tourism sector point of view. *Dogus University Dergisi*, 8, 36–46.
- Hanck, C., Arnold, M., Gerber, A., & Schmelzer, M. (2019). *Introduction to econometrics with R*. University of Duisburg-Essen.
- Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs*, 76(4), 408–420.
- Hernán, M. A., Hernández-Díaz, S., Werler, M. M., & Mitchell, A. A. (2002). Causal knowledge as a prerequisite for confounding evaluation: An application to birth defects epidemiology. *American Journal of Epidemiology*, 155(2), 176–184.
- Imai, K., & Kim, I. S. (2016). *When should we use linear fixed effects regression models for causal inference with panel data?* Princeton University.
- Judd, C. M., & Kenny, D. A. (1981). Process analysis: Estimating mediation in treatment evaluations. *Evaluation Review*, 5(5), 602–619.
- Kim, Y., & Steiner, P. M. (2021). Causal graphical views of fixed effects and random effects models. *British Journal of Mathematical and Statistical Psychology*, 74(2), 165–183.
- Lundberg, I., Johnson, R., & Stewart, B. M. (2021). What is your estimand? Defining the target quantity connects statistical evidence to theory. *American Sociological Review*, 86(3), 532–565.
- Maxwell, S. E., & Cole, D. A. (2007). Bias in cross-sectional analyses of longitudinal mediation. *Psychological Methods*, 12(1), 23–44.
- Maziarz, M. (2015). A review of the Granger-causality fallacy. *The Journal of Philosophical Economics: Reflections on Economic and Social Issues*, 8(2), 86–105.
- McNamee, R. (2003). Confounding and confounders. *Occupational and Environmental Medicine*, 60(3), 227–234.
- McNamee, R. (2005). Regression modelling and other methods to control confounding. *Occupational and Environmental Medicine*, 62(7), 500–506.
- Meehl, P. E. (1970). Nuisance variables and the ex post facto design. In M. Radner & S. Winokur (Eds.), *Minnesota studies in the philosophy of science* (Vol. 4, pp. 373–402). University of Minnesota Press.
- Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., Houts, R., Poulton, R., Roberts, B. W., Ross, S., Sears, M. R., Thomson, W. M., & Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences, USA*, 108(7), 2693–2698.
- Morabia, A. (2011). History of the modern epidemiological concept of confounding. *Journal of Epidemiology & Community Health*, 65(4), 297–300.
- National Center for Education Statistics. (2019). Young adult educational and employment outcomes by family socioeconomic status. *Condition of Education*. <https://nces.ed.gov/programs/coe/indicator/tbe>
- Pearl, J. (1995). Causal diagrams for empirical research. *Biometrika*, 82(4), 669–688.
- Pearl, J. (1998). Why there is no statistical test for confounding, why many think there is, and why they are almost right. *Escholarship*. <https://escholarship.org/uc/item/2hw5r3tm>



- Pearl, J. (2009). *Causality: Models, reasoning, and inference* (2nd ed.). Cambridge University Press.
- Pearl, J., Glymour, M., & Jewell, N. P. (2016). *Causal inference in statistics: A primer*. John Wiley & Sons.
- The Pell Institute. (2018). *Indicators of higher educational equity in the United States*.
- Peters, J., Janzing, D., & Schölkopf, B. (2017). *Elements of causal inference: Foundations and learning algorithms*. The MIT Press.
- Pourhoseingholi, M. A., Baghestani, A. R., & Vahedi, M. (2012). How to control confounding effects by statistical analysis. *Gastroenterology and Hepatology From Bed to Bench*, 5(2), 79–83.
- R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Roberts, B. W., & Wood, D. (2006). Personality development in the context of the neo-socioanalytic model of personality. In D. Mroczek & T. Little (Eds.), *Handbook of personality development* (pp. 11–39). Erlbaum.
- Roberts, B. W., Wood, D., & Smith, J. L. (2004). Evaluating Five Factor Theory and social investment perspectives on personality trait development. *Journal of Research in Personality*, 39(1), 166–184.
- Rohrer, J. (2019, April 16). Longitudinal data don't magically solve causal inference. *The 100% CI*. [the100.ci/2019/04/16/longitudinal-data](https://the100.ci/2019/04/16/longitudinal-data)
- Rohrer, J. M. (2018). Thinking clearly about correlations and causation: Graphical causal models for observational data. *Advances in Methods and Practices in Psychological Science*, 1(1), 27–42.
- Rohrer, J. M., Hünermund, P., Arslan, R. C., & Elson, M. (2021, February 1). *That's a lot to PROCESS! Pitfalls of popular path models*. PsyArXiv. <https://doi.org/10.31234/osf.io/paeb7>
- Ross, C. E., & Reskin, B. F. (1992). Education, control at work, and job satisfaction. *Social Science Research*, 21(2), 134–148.
- Savalei, V. (2019). A comparison of several approaches for controlling measurement error in small samples. *Psychological Methods*, 24(3), 352–370.
- Sechrest, L. (1963). Incremental validity: A recommendation. *Educational and Psychological Measurement*, 23(1), 153–158.
- Shear, B. R., & Zumbo, B. D. (2013). False positives in multiple regression: Unanticipated consequences of measurement error in the predictor variables. *Educational and Psychological Measurement*, 73, 733–756.
- Shmueli, G. (2010). To explain or to predict? *Statistical Science*, 25(3), 289–310.
- Simonsohn, U. (2019, November). Interaction effects need interaction controls. *Datacolada*. <http://datacolada.org/80>
- Slominski, L., Sameroff, A., Rosenblum, K., & Kasser, T. (2011). Longitudinal predictors of adult socioeconomic attainment: The roles of socioeconomic status, academic competence, and mental health. *Development and Psychopathology*, 23(1), 315–324. <https://doi.org/10.1017/S0954579410000829>
- Smith, G. T., Fischer, S., & Fister, S. M. (2003). Incremental validity principles in test construction. *Psychological Assessment*, 15(4), 467–477.
- Steiner, P. M., & Kim, Y. (2016). The mechanics of omitted variable bias: Bias amplification and cancellation of offsetting biases. *Journal of Causal Inference*, 4(2), Article 20160009.
- Sutin, A. R., Costa, P. T., Jr., Miech, R., & Eaton, W. W. (2009). Personality and career success: Concurrent and longitudinal relations. *European Journal of Personality*, 23(2), 71–84.
- U.S. Bureau of Labor Statistics. (2020). *Unemployment rates and earnings by educational attainment*. <https://www.bls.gov/emp/chart-unemployment-earnings-education.htm>
- VanderWeele, T. J. (2019). Principles of confounder selection. *European Journal of Epidemiology*, 34(3), 211–219.
- VanderWeele, T. J., & Shpitser, I. (2013). On the definition of a confounder. *Annals of Statistics*, 41(1), 196–220.
- Wang, Y. A., & Eastwick, P. W. (2020). Solutions to the problems of incremental validity testing in relationship science. *Personal Relationships*, 27(1), 156–175.
- Westfall, J., & Yarkoni, T. (2016). Statistically controlling for confounding constructs is harder than you think. *PLOS ONE*, 11(3), Article e0152719. <https://doi.org/10.1371/journal.pone.0152719>
- Westreich, D., & Greenland, S. (2013). The table 2 fallacy: Presenting and interpreting confounder and modifier coefficients. *American Journal of Epidemiology*, 177(4), 292–298.
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag. <https://ggplot2.tidyverse.org>
- Wilmot, M. P., & Ones, D. S. (2019). A century of research on conscientiousness at work. *Proceedings of the National Academy of Sciences, USA*, 116(46), 23004–23010.
- Wright, S. (1934). The method of path coefficients. *The Annals of Mathematical Statistics*, 5(3), 161–215.
- Yarkoni, T., & Westfall, J. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science*, 12(6), 1100–1122.