METHODOLOGY
European Journal of Research Methods for the Behavioral and Social Sciences
PsychOpen GOLD

Original Article

Check for updates

Revisiting Equivalent Structural Models Through the Mediation Models' Markov Equivalence Class

Dakota W. Cintron [1], Felix J. Thoemmes [2]

[1] *Division of Behavioral and Organizational Sciences, Claremont Graduate University, Claremont, CA, USA.*
[2] *Department of Psychology, Cornell University, Ithaca, NY, USA.*

Methodology, 2026, Vol. 22(3), 275–298, https://doi.org/10.5964/meth.17609

Received: 2025-04-07 • **Accepted:** 2026-04-30 • **Published (VoR):** 2026-08-31

Handling Editor: Jochen Mayerl, Chemnitz University of Technology, Chemnitz, Germany

Corresponding Author: Dakota W. Cintron, Division of Behavioral and Organizational Sciences, Claremont Graduate University, 150 E. 10th St., Claremont, CA 91711, USA. E-mail: dakota.cintron@cgu.edu

Abstract

Rational inquiry in psychology aims not just to describe phenomena but also to explain them. Researchers continue to use the classic trivariate mediation model as a key tool in this pursuit. The problem of equivalent structural models is revisited as a Markov equivalence class (MEC) (Andersson et al., 1997), illustrated using the trivariate mediation model. We highlight how the MEC complicates causal interpretation of the classic mediation model. To illustrate this, we examine two recent articles from a widely read psychology journal that apply the classic trivariate mediation model, one demonstrating strong causal control, the other weak causal control. We also introduce an R Shiny application that helps users understand this linkage and explore how causal assumptions shape the number of equivalent models in the MEC of a mediation model.

Keywords

equivalent models, causality, mediation, causal structure learning

Equivalent structural models are models with different structural representations that result in the same data-model fit (Hershberger & Marcoulides, 2013). More specifically, equivalent structural models are those that yield identical model-implied moments and goodness-of-fit indices when analyzing the same data. In the literature on structural equation modeling, considerable methodological attention has been paid to understanding the difficulties in inferring causal relations and substantive interpretations from equivalent structural models (Hershberger & Marcoulides, 2013; Lee & Hershberger, 1990; MacCallum et al., 1993; Raykov & Marcoulides, 2007; Stelzl, 1986). The existence of equivalent models poses serious challenges to the inferences typically made by researchers employing structural equation models (SEMs). Even if a hypothesized model fits the data better than competing models, it cannot be unequivocally confirmed if there are equivalent structural models. Although equivalent structural models based on different assumptions about the data (e.g., temporal precedence) can be ruled out, some level of doubt may always remain in such conclusions if equivalent models exist.

A related area of research that has received less attention in the SEM literature is the development of model equivalence arising from graph theory (Hershberger & Marcoulides, 2013). Graph theorists, in their attempts to formulate a mathematical basis for causal inference, have developed algorithms to *learn* equivalent causal model structures, where the algorithms vary in their generality (Heinze-Deml et al., 2018). Some of these *causal structure learning* algorithms (Drton & Maathuis, 2017; Heinze-Deml et al., 2018) are useful for recursive latent variable models (Spirtes et al., 2000) whereas others are suitable for non-recursive latent variables with feedback loops and correlated errors (Mooij & Claassen, 2020; Richardson, 1997). An important concept from this literature is the Markov Equivalence Class (MEC). The MEC is a collection of Directed Acyclic Graphs (DAGs) that encode the same set of conditional independencies. When attempting to estimate, infer, or learn a causal DAG from observational data, the exact underlying DAG is typically not identifiable. However, in general, one can obtain an MEC that contains all Markov equivalent DAGs (Heinze-Deml et al., 2018). The size of a MEC, that is, the number of DAGs it contains, is crucial to identifying possible causal models and understanding the range of potentially inferable causal effects.

In the remainder of this article, we formally define equivalent models and illustrate their challenges using the classic trivariate mediation model (Baron & Kenny, 1986), which remains widely used in psychology and is central to causal reasoning. We present an analytical example to demonstrate how issues of statistical conclusion validity can arise when inferring causality within this model. Next, we introduce causal structure learning algorithms, focusing on two fundamental methods, the PC and FCI algorithms, to identify the MEC of the trivariate mediation model (i.e., the set of acyclic graphs that imply the same pattern of conditional independence relations and are therefore observationally indistinguishable from the mediation model).[1] We then examine three key causal assumptions that determine which models are included in the MEC and demonstrate how these assumptions can reduce its size using a novel R Shiny application (Cintron & Thoemmes, 2026) available at https://dwcintron.shinyapps.io/mec-app/. Two published examples are used to illustrate cases of strong and weak causal control. We

1) Because the FCI algorithm allows for latent confounders, its output is what is known as a Partial Ancestral Graph (PAG), which represents a Markov equivalence class of Maximal Ancestral Graphs (MAGs) rather than DAGs. Throughout this article, we use 'MEC' loosely to refer to both the DAG-based equivalence class identified by the PC algorithm and the broader MAG equivalence class identified by the FCI algorithm, reserving more precise terminology where the distinction is consequential. See The FCI Algorithm section.

conclude by summarizing the main findings, discussing their implications, and outlining directions for future research.

A Formal Definition and Analytical Demonstration of Equivalent Structural Models

A typical SEM consists of the vector $\mathbf{x} = (X_1, X_2,..., X_p)'$ for p observed variables, a population covariance matrix $\mathbf{\Sigma}$, a sample covariance matrix $\mathbf{S}$, a vector θ consisting of parameters for the structural equation model M, and the population covariance matrix implied by M at θ, $\mathbf{\Sigma}(\widehat{\theta})$. Two alternative hypothesized models, Ml and M2, are considered equivalent if, for a given sample covariance matrix $\mathbf{S}$, the model-implied matrices are identical or, in other words, $\mathbf{\Sigma}\left(\widehat{\theta}\right)_{M1} = \mathbf{\Sigma}\left(\widehat{\theta}\right)_{M2}$ (Hershberger & Marcoulides, 2013).

Although we do not discuss model fit issues (Kenny & McCoach, 2003), since all equivalent models have the same model data fit, we briefly introduce the issue of model identification. Identification involves ensuring that all of the parameters in a model are uniquely identified. If the individual parameters in the model are uniquely identified, then the model as a whole is identified. With SEMs, the number of parameters that we can freely estimate is limited by the number of unique elements in the sample variance-covariance matrix $\mathbf{S}$. We cannot estimate more parameters than there are unique elements in the sample variance-covariance matrix $\mathbf{S}$ (Bollen, 1989).

Not all equivalent SEMs are identified, including some models within the trivariate mediation model MEC. Distinguishing which structural models are identified, unidentified, or outside the MEC is essential. If a model is not identified, its parameters cannot be uniquely estimated from the data without additional assumptions (e.g., fixed parameters, zero constraints, or exclusion restrictions). To evaluate how identification affects the size of the trivariate mediation MEC, our R Shiny application allows the user to toggle between identified versus unidentified models in the MEC. For instance, cyclical causal graphs are excluded because models in the trivariate mediation model's MEC are defined over acyclic graphs, and the feedback present in cyclical structures both violates acyclicity and generally leads to nonidentification absent further constraints.

It is important to consider unidentified models when evaluating possible causal graphs among three variables because causal claims about mediation are often made even when multiple observationally equivalent, yet unidentified, structures remain plausible. For example, a researcher may interpret a randomized X ⟶ M ⟶ Y design as supporting mediation, but alternative equivalent graphs involving reciprocal relations or unmeasured confounding between M and Y may be consistent with the observed covariance structure but not identified without further assumptions. Explicitly including unidentified models clarifies which causal conclusions depend on untestable identification assumptions and prevents overinterpretation of mediation effects. Thus, considering both identified and unidentified models provides a more complete representation of the causal ambiguity inherent in trivariate mediation analyses and highlights the additional assumptions required to uniquely support a given mediation graph.

An Analytical Example of Equivalent Models

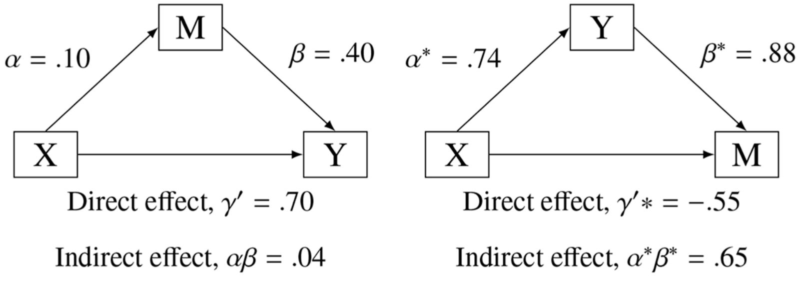
The classic trivariate mediation model consists of variables X, M, and Y. It is a causal model that assumes that the independent variable X causes Y and that X causes the mediator M which, in turn, causes the outcome variable Y. Causal inference with the trivariate mediation model is particularly challenging. For example, it is well known that a mediator in the trivariate X-M-Y mediation model cannot be statistically distinguished from a confounder (Fiedler et al., 2011; MacKinnon et al., 2000).

On occasion, one can still find the lingering belief in the literature that switching arrows in a single model and examining the resulting path coefficients (e.g., the indirect effect) can inform one about the veracity of any given model (Rohrer et al., 2022; Thoemmes, 2015; Thoemmes & Lemmer, 2019). Although these alternative models rarely make it to publications, it is not unusual that reviewers request that in addition to the mediation model presented, some arrows be reversed and the model tested again in order to determine a "winning" model. As Thoemmes (2015) writes: "The thought behind this practice is that if the mediated effect of one model is significant, but the other model has no significant mediated effect, then the mediation model with the significant indirect effect is the more plausible one" (p. 226). This belief is incorrect.

Based closely on the example of Thoemmes (2015), we also consider two hypothesized mediation models that might be portrayed in Figure 1. We assume that Model Ml (left model) is the true model and that Model M2 (right model) is incorrect and has the arrow between M and Y reversed.[2] For both models, we have the same sample correlation matrix (**S**):

Figure 1

Models Ml (left) and M2 (right)



2) We only consider standardized variables in this example, but our results also hold if unstandardized variables are used.

$$\mathbf{S} = \begin{bmatrix} 1.00 & & \\ 0.10 & 1.00 & \\ 0.74 & 0.47 & 1.00 \end{bmatrix}$$

The correlations of the three variables in the true model are given by: $\rho_{XM} = \alpha$, $\rho_{MY} = \beta + \alpha\gamma'$, and $\rho_{XY} = \gamma' + \alpha\beta$. These correlations can be quickly verified using Wright's tracing rules. Using the recursive formula from Cramér (1946), we can derive partial regression coefficients for Model M1. The partial regression coefficients for M1 are $\beta_{XM} = \rho_{XM} = \alpha$, $\beta_{MY \cdot X} = \frac{\rho_{MY} - \rho_{XM}\rho_{XY}}{(1 - \rho^2{}_{XM})} = \beta$, and $\beta_{XY \cdot M} = \frac{\rho_{XY} - \rho_{XM}\rho_{MY}}{(1 - \rho^2{}_{XM})} = \gamma'$. The set of parameter estimates $\left(\widehat{\theta}_{M1}\right)$ obtained from using the sample correlation matrix (**S**) with the partial regression coefficient formulas for M1 are (also shown in Figure 1):

$$\widehat{\theta}_{M1} = \begin{bmatrix} \alpha = & 0.10 \\ \beta = & 0.40 \\ \gamma' = & 0.70 \\ \alpha\beta = & 0.04 \end{bmatrix}$$

Again, using the recursive formula, we can derive the partial regression coefficients for Model M2. The partial regression coefficients for M2 are $\beta_{XY} = \rho_{XY} = \alpha^*$, $\beta_{YM \cdot X} = \frac{\rho_{MY} - \rho_{XM}\rho_{XY}}{(1 - \rho^2{}_{XY})} = \beta^*$, and $\beta_{XM \cdot Y} = \frac{\rho_{XM} - \rho_{MY}\rho_{XY}}{(1 - \rho^2{}_{XY})} = \gamma'^*$. Therefore, in the equivalent reversed arrow Model, M2, the set of parameter estimates $(\widehat{\theta}_{M2})$ obtained using the sample correlation matrix (**S**) and the partial regression coefficient formulas for M2 are (also shown in Figure 1):

$$\widehat{\theta}_{M2} = \begin{bmatrix} \alpha^* = & 0.74 \\ \beta^* = & 0.88 \\ \gamma'^* = & -0.55 \\ \alpha^*\beta^* = & 0.65 \end{bmatrix}$$

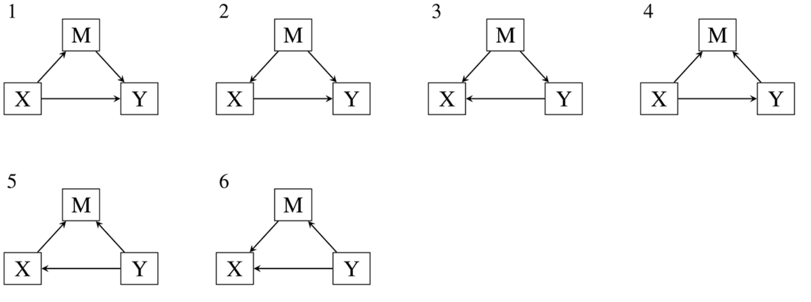
In the true model, M1, there is a strong direct effect of X on Y (γ' = .70), a weak effect of X on the mediator M (α = .10), and a medium effect of the mediator M on the outcome Y (β = .40). The indirect effect $\alpha\beta$ is very small at .04. In the reversed arrow model, the direct effect of X on M is large and negative (γ'^* = -.55), the effect of X on Y is large (α^* = .74), and the effect of Y on M is also large (β^* = .88). Consequently, the indirect effect is considerably larger ($\alpha^*\beta^*$ = .65) than in Model M1. Despite these vastly different parameter coefficients, the two models are indeed equivalent. Using any of the magnitudes of the path coefficients as a guide to model selection will fail, certainly if a heuristic for model selection would be to pick a model that shows large (indirect) effects.

The model-implied covariance matrices and fit indices from the model of M1 and M2 are likewise identical, despite the fact that the path coefficients are so dissimilar.

Beyond M1 and M2, many more equivalent mediation models can be estimated, each with different path coefficients but identical fit indices. Assuming that we allow for the presence of unobserved latent variables, there are 109 admissible acyclic graph configurations that are consistent with the observed data, of which 25 are statistically identified in the SEM sense and 84 are not identified without additional assumptions.[3] To see how we arrive at 25 identified DAGs, consider that we could form them by allowing three edge types between pairs of variables. These edge types are a forward directed arrow, a backward directed arrow, or a bidirected arrow.[4] Ruling out 2 DAGs with cycles, we are left with 25 identified DAGs (i.e., 3 × 3 × 3 − 2). Without the assumption of latent variables, and thereby forbidding the presence of bidirected arrows, there are only 6 identified DAGs in the trivariate mediation models' MEC (see Figure 2). Importantly, none of these identified DAGs can be distinguished on the basis of statistical significance or effect magnitude because they are observationally equivalent, implying the same model-implied covariance matrix and identical fit indices (MacCallum et al., 1993; Stelzl, 1986).

Figure 2

Six Identified Models in the Trivariate Mediation Model MEC (Assuming No Confounding)



In the classic trivariate mediation model, it may be relatively straightforward to draw all 25 identified DAGs in the MEC by hand. However, causal learning algorithms make this process much simpler and routinized, which can be particularly advantageous in more complicated models. In the next section, we introduce causal structure learning

3) See The FCI Algorithm section for how we arrive at 109 graph configurations. Raykov and Marcoulides (2007) highlight that there can be a potentially infinite number of equivalent models.

4) We use the bidirected arrow as a notional shorthand to represent that there is an unobserved latent variable that influences both connected variables (this is also referred to as latent projection; Richardson et al. 2023).

algorithms that allow us to learn the MEC among a set of variables (i.e., the set of equivalent models), with a particular set of causal constraints (e.g., no unmeasured confounding or temporal precedence), for a given set of variables and data. These algorithms can provide better insight into the assumptions underlying various equivalent structural models (e.g., allowing for unmeasured confounding or not) and lead to better theoretical design (Didelez, 2024). Moreover, causal learning algorithms have the flexibility to enumerate the broader set of 109 admissible graph configurations, spanning both the 25 identified and 84 unidentified structures, the latter of which would be invisible to approaches restricted to identified DAGs alone, making them particularly valuable for understanding the full scope of causal ambiguity in a given model.

Learning the Mediation Markov Equivalence Class of a Model

Given the apparent problem of equivalent models, we must consider all possible equivalent models for the trivariate mediation model. One way to discover all equivalent models for a given model is to *learn* the MEC of the model (Heinze-Deml et al., 2018). Based on the concept of learning in Bayesian networks (Chickering, 2002; Cintron, 2023), various algorithms for learning causal structure, each with distinct identification and causal assumptions, can estimate an MEC that represents the set of equivalent structural models for a given set of variables and data. Two popular algorithms that we focus on here are the PC and FCI algorithms (Spirtes et al., 2000).

Causal Structure Learning Algorithms

Causal structure learning is a data-driven model search approach that attempts to estimate or learn a causal graphical model that best describes the dependence structure among the observed data (Heinze-Deml et al., 2018). In general, a graphical model is said to be a representation of a multivariate distribution where a graph encodes conditional independence relationships among random variables (Lauritzen, 1996). A special type of graphical model is a causal graphical model in which the edges are interpreted as direct causal effects.

Important assumptions about the causal structural model for these learning algorithms include causal sufficiency, causal faithfulness, and acyclicity. *Causal Sufficiency* is the assumption that all relevant variables that influence the outcome and the exposure are included in the model. In other words, the assumption implies that there are no latent (or hidden) variables or no unmeasured confounding. *Causal Faithfulness* is the assumption that if a conditional independence relationship is observed in the data, there is a corresponding absence of a direct causal relationship between the variables involved. In other words, if two variables are conditionally independent given a set of other variables, there is no direct causal relationship between them. *Acyclicity* is the assumption that there are no directed cycles or loops in the graph. In other words, the graph is recursive.

Causal structure learning algorithms belong to one of five general classes: constraint-based methods, score-based methods, hybrid methods, SEMs-based methods, and methods exploiting invariance properties (Heinze-Deml et al., 2018). The two we focus on, the PC and FCI algorithms, are constraint-based methods. Constraint-based methods employ independence tests to identify a set of edge constraints for a graph and then find the acyclic graph structures that satisfy the constraints. Next, we detail the algorithms and present results from their implementation in Rpy-tetrad (Ramsey & Andrews, 2023) using the sample correlation matrix from the analytical example above. Importantly, these algorithms do not require the assumption of *Gaussianity* (i.e., the assumption that the errors are normally distributed). They can be implemented with distribution-free conditional independence tests. In practice, however, linear-Gaussian assumptions are often adopted in SEM applications, which leads to equivalence classes of models that share the same implied covariance structure.

Our discussion of indistinguishability among mediation models is therefore situated within this conventional linear SEM framework. Under these assumptions, multiple causal structures can be observationally equivalent because they imply the same covariance matrix. If one is willing to impose additional assumptions in the distributional or functional form, such as non-Gaussian errors as in LiNGAM (Shimizu et al., 2006, 2011), the causal directions may become identifiable and the size of the equivalence class can be reduced. We do not pursue these approaches here; instead, our focus is on how qualitative causal assumptions (e.g., confounding, temporal ordering, and identification constraints) shape the set of admissible mediation graphs.

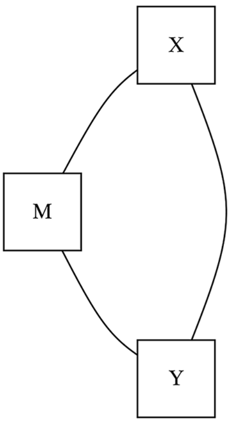
The PC Algorithm

The PC (Peter-Clark) algorithm is one of the oldest structure learning algorithms (Glymour & Zhang, 2019; Spirtes et al., 2000). The PC algorithm is named after the developers of the algorithm, Peter Spirtes and Clark Glymour. The primary assumptions of the PC algorithm include acyclicity, causal faithfulness, and causal sufficiency. To learn the structure of the underlying DAG, the algorithm performs numerous conditional independence tests. Specifically, the PC algorithm learns the completed partially directed acyclic graph (CPDAG) (Andersson et al., 1997; Chickering, 2002) of the underlying DAG in three steps: 1) identify the skeleton, 2) determine the v-structures, and 3) identify edge orientations. Note, a v-structure in a DAG is a specific configuration of nodes and edges. In a v-structure, two parent nodes (e.g., A and B) converge into a single child node (C), but are not directly connected to each other. A CPDAG can have one of three edge orientations $i \longrightarrow j$, $i \longleftarrow j$, $i - j$ with the latter undirected edge indicating uncertainty about the direction of the arrow.

The result of the PC algorithm for the sample correlation matrix (**S**) previously described is given in Figure 3. The results indicate that both $\longleftarrow$ and $\longrightarrow$ relationships are allowed between each of the variables in the model (i.e., this means that no causal direction is determined, but all DAGs that include either of these edges are equivalent). Consequently, 6 potential DAGs are part of the MEC learned by the PC algorithm (i.e., 2 possible edges for each of the relationships between X-M, X-Y, and M-Y minus those that result in cycles or $2 \times 2 \times 2 - 2 = 6$). Note that there are only 6 DAGs in the mediation MEC learned by the PC algorithm because it does not allow for the presence of latent or hidden variables between nodes and it does not allow for models with cycles. These results can be verified using the R Shiny application. As will be shown in the next section, a major difference with the FCI algorithm compared to the PC algorithm is that the FCI algorithm allows for the presence of latent or hidden variables between nodes when learning the structure of the graph (i.e., the algorithm relaxes the causal sufficiency assumption).

Figure 3

*Result of the PC Algorithm for the Sample Correlation Matrix **S***



The FCI Algorithm

The FCI (Fast Causal Inference) algorithm is a modification of the PC algorithm that no longer requires the assumption of causal sufficiency by allowing arbitrarily many latent variables. Unlike the PC algorithm, which outputs a CPDAG representing a DAG-based MEC, the output of the FCI algorithm is a partial ancestral graph (PAG), a compact representation of a Markov equivalence class of Maximal Ancestral Graphs (MAGs). MAGs extend DAGs by permitting bidirected edges that encode the presence of latent confounders, making the FCI-based equivalence class broader than the DAG-based MEC identified by the PC algorithm. A PAG contains the edges $i \longrightarrow j, i \longleftarrow j, i \leftrightarrow j, i \circ\!\!-\!\!> j, i <\!\!-\!\!\circ\ j, i \circ\!\!-\!\!\circ\ j$.

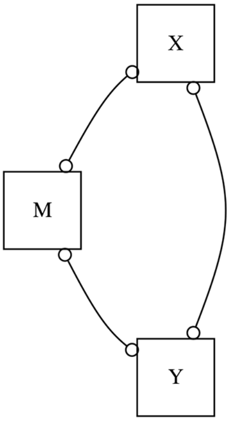
The interpretation of the edge marks is as follows. An edge mark pointing in a given direction indicates that this edge type is present in all MAGs in the equivalence class represented by the PAG, whereas the circle mark indicates uncertainty about the directionality of an edge. Consequently, if both variables in a bivariate relation are marked with circles (i.e., the edge between i and j is $\circ\!\!-\!\!\circ$), the following five edge types are possible (i.e., $i \longrightarrow j, i \longleftarrow j, i \leftrightarrow j, i \underset{\longrightarrow}{\leftrightarrow} j$, or $i \underset{\longleftarrow}{\leftrightarrow} j$) A bidirected edge between two variables indicates that there is at least one unmeasured confounder (or, equivalently, latent variable) between the variables (Glymour & Zhang, 2019). The last two edge types allow for directional knowledge in the presence of potential unmeasured confounding or latent variables. That is, we have either an edge that has a forward pointing arrow and a bidirected edge or a backward pointing arrow and a bidirected edge. If an edge type has a circle and directed edges, e.g., $i \circ\!\!- > j$, then the following edge types are possible (i.e., $i \longrightarrow j$, $i \leftrightarrow j$, or ($i \underset{\longrightarrow}{\leftrightarrow} j$). In the case of $i < -\!\!\circ\, j$, the directed arrows would be reversed.

The steps of the FCI algorithm are similar to the steps of the PC algorithm (Heinze-Deml et al., 2018). However, the FCI algorithm conducts additional tests to learn the correct skeleton and orient v-structures in the presence of latent or hidden variables. The result of the FCI algorithm, for the sample correlation matrix (**S**) above, without causal constraints, is given in Figure 4. The results indicate that for each variable pair in the model, all five edge types remain possible, reflecting maximal uncertainty about causal direction and the presence of latent confounding. Accordingly, in the trivariate mediation setting, there are 125 possible combinations of edge types across the three variable pairs (X–M, X–Y, and M–Y) when allowing five edge types per pair (i.e., 5 × 5 × 5 = 125). Of the 125 possible combinations, 16 contain directed cycles and are therefore excluded, yielding 109 admissible acyclic graphs, a result that can be verified using the R Shiny application.[5]

5) The 16 excluded combinations arise from two possible directed cycles among three variables: a clockwise cycle ($X \longrightarrow M \longrightarrow Y \longrightarrow X$) and a counterclockwise cycle ($X \longrightarrow Y \longrightarrow M \longrightarrow X$). Each cycle requires directed edges pointing in the cycle direction across all three variable pairs. Since edge types $i \longrightarrow j$ and $i \underset{\longrightarrow}{\leftrightarrow} j$ both carry a forward directed component, and $i \longleftarrow j$ and $i \underset{\longleftarrow}{\leftrightarrow} j$ both carry a backward directed component, each cycle yields 2 × 2 × 2 = 8 combinations, giving 8 + 8 = 16 total cycle-containing combinations with no overlap between the two directions.

Figure 4

Result of the FCI Algorithm for the Sample Correlation Matrix $\boldsymbol{S}$



Reducing the MEC of the Trivariate Mediation Model: The Role of Causal Reasoning and Control

Given that equivalent models of the mediation model cannot be ruled out using statistical tests, a relevant question may then be whether it is possible to reduce the size of the equivalence class. One way that a researcher can reduce the size of the MEC and narrow in on a particular model or set of equivalent models is by invoking causal assumptions about the variables in the model. These causal assumptions could, for example, be the assumption that a particular relationship is not confounded (plausible under randomization, implausible in observational studies) or the temporal and causal precedence of one variable before the other.

To aid in our understanding of reducing the MEC of the popular trivariate mediation model, we have developed an R Shiny application (https://dwcintron.shinyapps.io/mec-app/) that displays the identified and unidentified graphs in the MEC under various causal assumptions. The application also displays models not in the trivariate mediation MEC by assuming models with missing edges (i.e., overidentified models with positive degrees of freedom due to a constraint of independence between variables) and models with directed cycles or feedback loops. The application allows the user to impose causal assumptions and thus reduce or expand the size of models in and out of the MEC. The application accepts several causal assumptions as inputs (e.g., whether X has been randomized or whether there is unmeasured confounding). Furthermore, the application includes a tab with instructions for using

TETRAD in R with the Rpy-tetrad package (Ramsey & Andrews, 2023), including worked examples of imposing causal assumptions such as randomization and temporal ordering.

Once the user has provided specific assumptions, the models that are no longer plausible given the assumption(s) are eliminated, and the display is updated to show the remaining models satisfying the assumptions. The absence of additional, unobserved confounding (beyond the bidirected arrows in the graphs) is not assumed but is available for exploration in the R Shiny application. This practically means that a researcher who has eliminated some models from the equivalence class may still not be able to estimate the path coefficients of these models without bias. It is only possible to rule out alternative models, but not to prove that the path coefficients of the remaining models are unbiased and thus "true."

Model Identification: Is the Model Identified?

We include the option to switch between identified and unidentified models because it is important to distinguish which causal structures are not only part of the MEC of the trivariate mediation model, but are also statistically identifiable and therefore estimable from the observed data. Although not a causal assumption, this assumption is critical for obtaining parameter estimates from a model with observed data. Identified models are those for which we can obtain a unique solution for every parameter. One necessary (though not sufficient) rule for an identified model is that it has nonnegative degrees of freedom (i.e., the number of estimated parameters is less than or equal to the number of elements in the sample covariance matrix). Although other sufficient identification conditions exist, such as the recursive rule described by Bollen (1989), these are beyond the scope of this review.

In total, there are 216 potential models (causal graphs) that we could assume from three variables and six different edge types (i.e., 6 × 6 × 6 = 216). The six edge types are no edge, $i \longrightarrow j$, $i \longleftarrow j$, $i \leftrightarrow j$, $i \underset{\longrightarrow}{\leftrightarrow} j$, or $i \underset{\longleftarrow}{\leftrightarrow} j$.[6] A bidirected edge with a second edge, either pointing forward or backward, signifies the case where we have knowledge of a directional relationship between two variables, but we still permit that unmeasured confounding or latent variables may bias the relationship. The 216 models range from the model without edges between variables (i.e., the empty graph) to directed paths between all variables in addition to unmeasured confounding between all variables. However, once we eliminate unidentified models, there are only 62 possible models that are identified and which could actually be estimated. Some of these models include

6) This reflects the app's default settings: (a) *Is the graph identified?* = *No*; (b) *Is the graph in the mediation MEC?* = *No*; (c) each *Is there confounding between…* option = *Yes* (i.e., confounding allowed); and (d) all temporal ordering options (e.g., *Is X causally/temporally before M?, Is M causally/temporally before Y?*) are set to *No*. Under these defaults, the application displays all 216 causal graphs, including identified and unidentified models both inside and outside the trivariate mediation MEC.

positive degrees of freedom (e.g., missing edges between variables) as well as those that are just-identified (i.e., there are zero degrees of freedom). Importantly, not all of these 62 identified models are in the trivariate mediation model MEC, which assumes no absent edges or directed cycles.

MEC Model Inclusion: Is the Model in the MEC of the Trivariate Mediation Model?

We include the option to switch between DAGs in a MEC because it is important to know which models are part of the classic trivariate mediation model's MEC. To be included in the trivariate mediation MEC, the structural model must not have missing edges. Although models with missing edges may be identified, they are not in the trivariate mediation MEC. That is, for the model to be in the trivariate mediation model MEC there must be at least one edge type between all variables. All of the models in the trivariate mediation MEC, therefore, are just-identified or unidentified. After filtering out unidentified models and models with missing edges, we are left with 25 identified DAGs in the trivariate mediation MEC. Recall from the discussion of the FCI algorithm, there are 109 acyclic graph configurations associated with the trivariate mediation model, of which 84 are unidentified structures.

Causal Assumption 1: Is There Confounding Between Pairs of Variables in the Trivariate Mediation Model?

Confounding is perhaps the most consequential threat to causal inference in mediation models, as unmeasured common causes can bias estimated effects along every pathway in the model. Assuming confounding or not allows for the presence of bidirected edges between pairs of variables or not. Considering potential confounding in mediation models is important for accurate estimation of effects and valid causal inference. Confounders can bias the relationships between variables, leading to incorrect conclusions about causal pathways. Including confounders ensures the validity and reliability of the results, providing a more accurate representation of the relationship of variables. In general, addressing confounding is crucial for obtaining meaningful and trustworthy insights from mediation analysis.

For each of the three variable pairs in the MEC of the trivariate mediation model, we may assume the absence or presence of unobserved confounding. The presence of an unmeasured, latent confounder between two variables would imply that we draw a bidirected arrow between two variables. This is usually a plausible assumption. A more stringent assumption (which is often difficult to defend) is to assume the absence of such unobserved, latent common causes.

Depending on which assumptions we are willing to make with respect to unobserved confounding, the MEC will change accordingly. If we could only assume no confounding for one of the pairs of variables, we would be left with 16 identified DAGs in the mediation MEC. If we could only assume no confounding for two of the pairs of variables, we would be left with 10 identified DAGs. Lastly, if we were to assume the absence of any confounding between the variables, we would be left with 6 identified DAGs in the mediation MEC. This is the same result as the PC algorithm.

Causal Assumption 2: What Are the Temporal Relations Among Variables?

Establishing the temporal order of variables is a foundational step in causal reasoning, and without it the direction of any edge in the graph remains ambiguous by default. Assuming temporal precedence among variables in a model allows for an edge to be directed between the variables. Temporal precedence ensures that the cause precedes the effect, which is fundamental for valid causal inference. Without considering the temporal sequence, the mediation analysis might incorrectly attribute causality, leading to misleading conclusions. This understanding improves the reliability and validity of the mediation model, providing more meaningful insights into the causal mechanisms at play.

A researcher may assume that there is a temporal and causal precedence from one variable to another. The existence of such temporal precedence implies that edges emanating from the temporally preceding variable must be either forward or bidirected, but cannot point back toward the temporally preceding variable. As an example, if we assume that X is temporally before M and assume the presence of unmeasured confounding between all variables, we reduce the MEC to 17 identified DAGs. Moreover, if we assume that X is also temporally before Y, we further reduce the MEC to 12 identified DAGs. Finally, if we further assume that M is temporally before Y, we are left with 8 identified DAGs in the mediation MEC. If we further eliminate confounding between all pairs of variables, we would be left with 1 DAG (i.e., the traditional mediation model).

Causal Assumption 3: Is X Randomized?

Among all the assumptions a researcher can invoke, randomization of X is uniquely powerful because it fulfills multiple causal assumptions simultaneously by design. Randomizing X guarantees that all DAGs that have arrows pointing into X can be ruled out. It also guarantees no unobserved common causes between X and any other variable in the model, though confounding between non-randomized variables such as M and Y may remain. By randomly assigning participants to different levels of X, researchers can ensure that any observed effects on M and Y are due to X, rather than other factors. This process enhances the internal validity of the study and eliminates selection bias, making the results more reliable and replicable.

Assuming that we are examining identified models in the trivariate mediation MEC, there are 3 possible DAGs remaining after randomizing X: the DAG with a forward arrow between M and Y, the DAG with a backward arrow between M and Y, and the DAG with a bidirected edge between M and Y. If we assume the temporal order between M and Y, we are left with 2 potential DAGs. Alternatively, if we rule out confounding between M and Y by assumption, we are left with 2 potential DAGs. These 2 DAGs are those of the analytical example in Figure 1.

Two Examples of Causal Control

In this section, we will provide two examples that show how one can use causal assumptions to minimize the size of the total class of equivalent models. In the context of both applied examples, we will discuss how tenable the causal assumptions are, and remind the reader that these assumptions are typically untestable and must be argued for based on either the design of a study or theoretical knowledge of the subject domain.

Strong Causal Control

Our first example is based on a published study by Shuman et al. (2021). The authors tested the effectiveness of nonviolent collective action in an effort to induce social change. We used this study as an example because it was recently published in one of the flagship journals on social psychology *(Journal of Personality and Social Psychology).* Among a set of a large number of studies, the authors examined the effects of different types of protest on the willingness of the opposing party to make concessions. Some of the studies were conducted using vignettes in which participants were randomly assigned to read a description of violent or nonviolent protest, and then mediators and outcomes were measured subsequently.

The specific study that we will use as an example was analyzed with a moderated mediation model. For simplicity, we will simplify this model by omitting the moderation effect and instead will only consider a mediation model for one subgroup. The study under consideration randomly assigned participants to read articles about different types of protest (e.g., normative nonviolent protest, nonnormative nonviolent protest, violent protest), then measured opinions about how disruptive participants judged the particular type of protest (operationalized through a measure called “constructive disruption”), and finally assessed an outcome measure, which was the support for concessions for the protesting group. The authors reported a mediation model in which the effect of the type of protest on the support for concessions was mediated by the amount of constructive disruption.

Without making any assumptions about the presence or absence of confounding, time, and causal ordering, or taking into account randomization, all possible models in the mediation MEC are equally supported by the data that were collected. That is to say that 25 identified models could be considered including the one that was presented in the paper. If we also consider unidentified models (i.e., models that could be the true data-generating models but cannot be estimated with the data at hand), we would have to consider 109 graph configurations that remain structurally plausible given the observed covariance structure, though many cannot be uniquely estimated without additional assumptions.

However, as we know from this study, there are in fact a number of assumptions that are plausible. First, the treatment (type of protest) was randomized. This information alone reduces the set of identified and unidentified models in the trivariate mediation MEC dramatically. We can exclude any model that has any arrow pointing into the treatment variable, and we further know that the temporal and causal ordering of the treatment is prior to both the mediator and the outcome. Based on the single design choice of randomization, we can exclude an astonishing 104 identified and unidentified graph configurations in the trivariate mediation model MEC, which leaves us only with five remaining graph configurations (three of them identified, and the remaining two models unidentified). This reduction in equivalent models truly highlights the power of randomization.

However, these remaining models are still indistinguishable from each other unless more causal assumptions are invoked. That is to say that the endorsed model in the article is just as well supported by the data as the remaining four models in the mediation MEC, once randomization of the treatment has reduced the number of models. We consider first assumptions about the causal and temporal ordering of the variables. Due to randomization, we already know that the treatment must be causally and temporally prior to both the mediator and the outcome. But what can we say about the ordering of the two variables that were not randomized, namely the mediator and the outcome?

A somewhat reasonable claim could be made about the causal ordering of these two variables. The mediator asked the participants about their perceived level of constructive disruption of the protest. It seems implausible that the willingness to make concessions would cause the perception of disruption. It is far more plausible that it was the type of protest that caused the perception of disruption. After all, the vignettes described the type of protest and included a description of the disruption that was caused. A potential counter-argument might be that participants would have general feelings towards a protest and willingness to concessions, which could influence the way that they perceive the disruption. For example, if participants were generally less favorable toward violent protests, and thus have diminished willingness for concessions for such protests, these feelings might influence their perception of the amount of disruption. Ultimately, this argument must be made on a theoretical basis, but if we are willing to accept the first argument, we may claim that the mediator is indeed causally prior to the outcome, and this will in fact allow us to reduce the MEC even further. Accepting this assumption leaves a total of only three equivalent models, and only two of them that are identified.

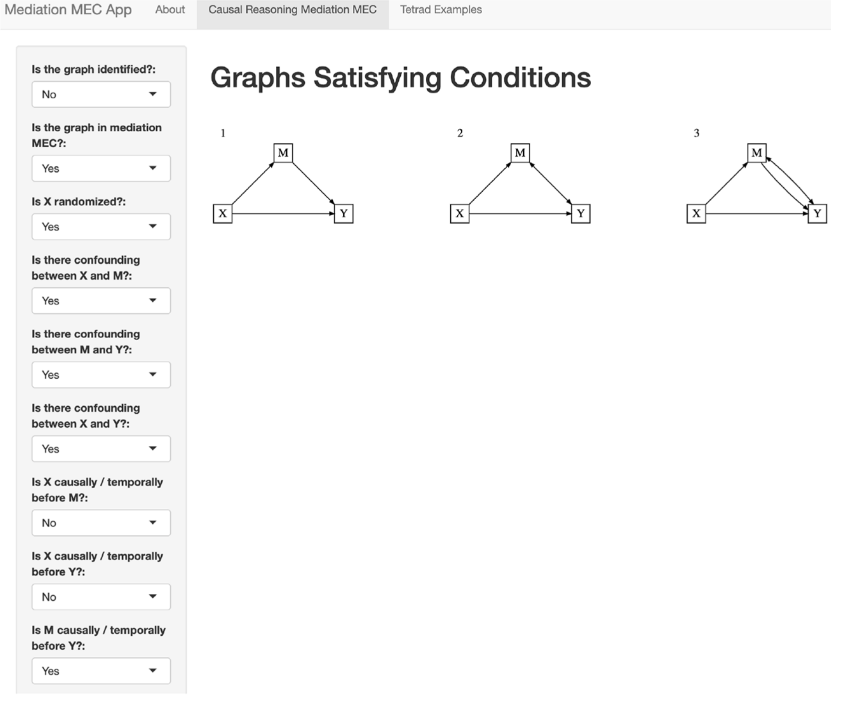
The remaining equivalent models (shown in Figure 5) all share that the randomized treatment (protest type) has a causal effect (i.e., directed arrows) on the mediator and the outcome. These paths in the mediation model are unconfounded (i.e., no bidirected arrows). Less clarity exists with respect to the relationship between the mediator and the outcome. In the three equivalent models, the relationship is either an unconfounded direct path from the mediator to the outcome, a purely confounded relationship (a bidirected arrow) between the mediator and the outcome, or a combination of an actual causal effect of the mediator on the outcome and a simultaneous confounding effect (i.e., directed arrow and bidirected arrow). This last model that includes two arrows between M and Y is not identified and cannot be estimated from data, without making additional assumptions about the strength of some of the coefficients.[7]

If we further assume that the researchers believed that there should be a directed path between the mediator and the outcome, and that the association between M and Y is not purely due to unobserved confounding, then only two equivalent models remain. The classic trivariate mediation model that was also published by the authors, or the same mediation model that in addition also includes a bidirected arrow between M and Y. This massive reduction in equivalent models (from 109 to 2) is possible through invoking what are relatively plausible assumptions in this context. The choice to favor the model without the bidirected arrow is based on an implicit assumption that the model with the bidirected arrow can be ruled out, or in other words, that the strength of any potential bidirected arrow is exactly zero. This assumption of zero confounding is often implausible in many instances of mediation models that only randomized X. Given that neither the mediator nor the outcome was randomized, and no other covariates were used for adjustment in the model, it seems implausible to assume that there is no confounding between the mediator and the outcome, making it impossible to rule out models in which there are bidirected arrows between the mediator and the outcome.

7) This lack of identification arises because the directed path from M to Y and the bidirected edge (representing unmeasured confounding) are not separately estimable from the observed covariance structure. Identification can be achieved only by imposing additional assumptions, such as ruling out unmeasured confounding (i.e., setting the bidirected edge to zero) or fixing the magnitude of the confounding effect and conducting sensitivity analyses over a range of plausible values (see, e.g., Alvarez-Bartolo & MacKinnon, 2025).

Figure 5

Shiny App Output Under Randomized Treatment and Mediator-Outcome Temporal Ordering



The final causal assumptions are based on causal and temporal ordering. The mediator and the outcome were collected as part of a single online questionnaire in this study, so it is difficult to make an argument in favor of strong temporal ordering.

Weak Causal Control

In the previous example, the equivalence class could be drastically reduced through a set of relatively plausible assumptions. We now present a second example in which additional equivalence classes may remain. In a paper by Ruisch et al. (2021), also published in the *Journal of Personality and Social Psychology,* researchers measured the density of fungiform papilla on the tongues of human subjects, then measured their sensitivity to disgust (through a questionnaire), and finally their political conservatism. Although the authors correctly warned against a causal interpretation of this model, we may still want to consider how many equivalent models would have been equally supported by the data. The assumed mediation model was that the density of papilla has an effect on disgust sensitivity, and that, in turn, has an effect on political conservatism. As in the previous study, we may start without any causal assumptions and consider that all 109 models, identified and unidentified, are equally supported. How could we narrow the equivalence class in this study?

First, papilla density is (obviously) not randomized, which means that a large number of equivalent models cannot be ruled out, and in fact, we are still left with the total 109 models. What other assumptions could plausibly be made about the causal and temporal ordering, and unconfoundedness between individual paths? First, it may seem plausible to assume that papilla density is both temporally and causally prior to both disgust sensitivity and political conservatism. Presumably, the density of taste buds on your tongue is genetically determined and cannot change based on political beliefs or disgust. However, it turns out that even this assumption is not supported by the current science on papilla density. As shown in a large-scale longitudinal epidemiological study, the Beaver Dam Offspring Study, papilla density is only 40% due to genetic factors, and the remainder is due to lifestyle choices, e.g., drinking or smoking behavior (Fischer et al., 2013). These types of behaviors could reasonably be caused or correlated with political affiliations or measured disgust. This implies that we cannot reasonably rule out any temporal ordering or lack of confounding between papilla density and the mediator and outcome variable. This again leaves us with the full set of 109 models.

The temporal and causal order between the mediator and outcome, conservatism, and political disgust, respectively, is even more difficult to argue. Given that the scale on disgust also contains items that, e.g., tap into feelings about sex education, it seems not entirely implausible that political conservatism could influence scores on the disgust scale, but likewise it is not implausible if the preferred causal ordering of the authors from disgust sensitivity to political conservatism is the correct one. Given this indeterminacy, it seems wise not to exclude any model from the equivalence class with respect to the causal and temporal ordering of the mediator and the outcome.

Lastly, in terms of unconfoundedness assumptions, it seems safest to assume that there is potential for confounding between all three involved variables, and that no bidirected arrow can be safely ruled out. This means that not a single one of the 109 models can be ruled out, and it seems highly suspect that the single model that was presented in the published research is in fact the correct one.

Discussion

This article presents an introduction to the problem of equivalent structural models for the trivariate mediation model using the notion of an MEC. We highlight the challenges of causal reasoning with models in an MEC. The exploration of the trivariate mediation model is particularly relevant due to its widespread application in psychological research and the critical role of causal inferences drawn from it. We highlight the potential pitfalls of relying solely on statistical significance to differentiate between equivalent models, which can often lead to misleading conclusions. We further clarify this issue using two case examples of published work in a leading psychology journal: one with nearly complete causal control and another with very little room for causal control.

Moreover, by developing an R Shiny application, we provide a valuable tool for researchers to visualize and enumerate the various structural models associated with the trivariate mediation models' MEC. This application not only aids in understanding the impact of causal assumptions on the size of the MEC but also serves as an educational resource for those looking to deepen their understanding of causal structure learning algorithms. The discussion of key causal assumptions and their implications for reducing the size of the MEC demonstrates the importance of careful consideration and application of causal theory in empirical research. The latter demonstration is in line with the Hershberger and Marcoulides (2013) call for researchers to enumerate how many models are equivalent to the hypothesized model and to understand the available preventive options before data collection begins. We encourage a more nuanced approach to model selection and interpretation, advocating for a balance between statistical evidence and theoretical plausibility.

Furthermore, this article provides examples of the utility of causal learning algorithms. Causal learning algorithms can be a useful tool for learning the MEC of a given model under a variety of modeling assumptions (e.g., cycles or confounding). The implications of causal learning algorithms extend beyond just learning the MEC of a model. For example, an interesting development in this space is the development of methods to identify causal effects and the bounds on causal effects given only observational data by utilizing the learnable MEC of a causal diagram (Bellot, 2023; Jaber et al., 2022).

In conclusion, this article contributes to the ongoing discourse on equivalent structural models, offering both theoretical insights and practical tools to advance the field of structural equation modeling. We highlight the utility of the MEC, causal structure learning algorithms, and causal reasoning and control for understanding the relations among variables in the trivariate mediation model using findings from empirical psychological research. Future research directions may include the extension of these methods to more complex models (e.g., longitudinal mediation) and the exploration of alternative algorithms for causal structure learning.

Funding: The authors have no funding to report.

Acknowledgments: The authors have no additional (i.e., non-financial) support to report.

Competing Interests: The authors have declared that no competing interests exist.

Data Availability: A Shiny App is openly available in the Supplementary Materials at Cintron and Thoemmes (2026).

Supplementary Materials

Type of supplementary material	Availability/Access
Data	
No data was made available	–
Code	
No code was made available	–
Material	
No supplementary material provided	–
Study/Analysis preregistration	
No preregistration	–
Other	
Shiny App Version 1.0	Cintron and Thoemmes (2026)

References

Alvarez-Bartolo, D., & MacKinnon, D. P. (2025). Three sensitivity-analysis methods to assess unmeasured pretreatment confounding bias in experimental mediation analysis. *Advances in Methods and Practices in Psychological Science, 8*(3). https://doi.org/10.1177/25152459251355586

Andersson, S. A., Madigan, D., & Perlman, M. D. (1997). A characterization of Markov equivalence classes for acyclic digraphs. *The Annals of Statistics, 25*(2), 505–541. https://doi.org/10.1214/aos/1031833662

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. https://doi.org/10.1037/0022-3514.5L6.1173

Bellot, A. (2023). *Towards bounding causal effects under Markov equivalence* [arXiv:2311.07259]. arXiv. https://arxiv.org/abs/2311.07259

Bollen, K. A. (1989). *Structural equations with latent variables.* John Wiley & Sons.

Chickering, D. M. (2002). Learning equivalence classes of Bayesian-network structures. *The Journal of Machine Learning Research, 2*, 445–498. https://jmlr.org/papers/v2/chickering02a.html

Cintron, D. W. (2023). Bayesian networks. In *International Encyclopedia of Education* (4th ed., pp. 119–129). Elsevier.

Cintron, D. W., & Thoemmes, F. J. (2026). *Mediation Markov equivalence class app* [Shiny Application Version 1.0]. https://dwcintron.shinyapps.io/mec-app/

Cramér, H. (1946). *Mathematical methods of statistics.* Princeton University Press.

Didelez, V. (2024). Invited commentary: Where do the causal DAGs come from? *American Journal of Epidemiology, 193*(8), 1075–1078. https://doi.org/10.1093/aje/kwae028

Drton, M., & Maathuis, M. H. (2017). Structure learning in graphical modeling. *Annual Review of Statistics and Its Application, 4*, 365–393. https://doi.org/10.1146/annurev-statistics-060116-053803

Fiedler, K., Schott, M., & Meiser, T. (2011). What mediation analysis can (not) do. *Journal of Experimental Social Psychology, 47*(6), 1231–1236. https://doi.org/10.1016/j.jesp.2011.05.007

Fischer, M. E., Cruickshanks, K. J., Schubert, C. R., Pinto, A., Klein, R., Pankratz, N., Pankow, J. S., & Huang, G.-H. (2013). Factors related to fungiform papillae density: The Beaver Dam Offspring Study. *Chemical Senses, 38*(8), 669–677. https://doi.org/10.1093/chemse/bjt033

Glymour, C., & Zhang, K. (2019). Review of causal discovery methods based on graphical models. *Frontiers in Genetics, 10*, Article 524. https://doi.org/10.3389/fgene.2019.00524

Heinze-Deml, C., Maathuis, M. H., & Meinshausen, N. (2018). Causal structure learning. *Annual Review of Statistics and Its Application, 5*, 371–391. https://doi.org/10.1146/annurev-statistics-031017-100630

Hershberger, S. L., & Marcoulides, G. A. (2013). The problem of equivalent structural models. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (2nd ed., pp. 73–112). Information Age Publishing.

Jaber, A., Ribeiro, A., Zhang, J., & Bareinboim, E. (2022). Causal identification under Markov equivalence: Calculus, algorithm, and completeness. *Advances in Neural Information Processing Systems, 35*, 3679–3690. https://proceedings.neurips.cc/paper_files/paper/2022/hash/17a9ab4190289f0el504bbb98dldllla-Abstract-Conference.html

Kenny, D. A., & McCoach, D. B. (2003). Effect of the number of variables on measures of fit in structural equation modeling. *Structural Equation Modeling, 10*(3), 333–351. https://doi.org/10.1207/sl5328007seml003_l

Lauritzen, S. L. (1996). *Graphical models (Vol. 17)*. Clarendon Press.

Lee, S., & Hershberger, S. (1990). A simple rule for generating equivalent models in covariance structure modeling. *Multivariate Behavioral Research, 25*(3), 313–334. https://doi.org/10.1207/sl5327906mbr2503_4

MacCallum, R. C., Wegener, D. T., Uchino, B. N., & Fabrigar, L. R. (1993). The problem of equivalent models in applications of covariance structure analysis. *Psychological Bulletin, 114*(1), 185–199. https://doi.org/10.1037/0033-2909.114.L185

MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention Science, 1*, 173–181. https://doi.org/10.1023/a:1026595011371

Mooij, J. M., & Claassen, T. (2020). Constraint-based causal discovery using partial ancestral graphs in the presence of cycles. *Proceedings of the 36th Conference on Uncertainty in Artificial Intelligence (UAI), 124*, 1159–1168. http://proceedings.mlr.press/vl24/m-mooij20a.html

Ramsey, J., & Andrews, B. (2023). Py-Tetrad and RPy-Tetrad: A new Python interface with R support for Tetrad causal search. *Proceedings of the 2023 Causal Analysis Workshop Series, 223*, 40–51. https://proceedings.mlr.press/v223/ramsey23a.html

Raykov, T., & Marcoulides, G. A. (2007). Equivalent structural equation models: A challenge and responsibility. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(4), 695–700. https://doi.org/10.1080/10705510701303798

Richardson, T. S. (1997). Extensions of undirected and acyclic, directed graphical models. *Proceedings of the Sixth International Workshop on Artificial Intelligence and Statistics, R1*, 407–420. https://proceedings.mlr.press/rl/richardson97a.html

Richardson, T. S., Evans, R. J., Robins, J. M., & Shpitser, I. (2023). Nested Markov properties for acyclic directed mixed graphs. *The Annals of Statistics, 51*(1), 334–361. https://doi.org/10.1214/22-aos2253

Rohrer, J. M., Hünermund, P., Arslan, R. C., & Elson, M. (2022). That's a lot to process! Pitfalls of popular path models. *Advances in Methods and Practices in Psychological Science, 5*(2). https://doi.org/10.1177/25152459221095827

Ruisch, B. C., Anderson, R. A., Inbar, Y., & Pizarro, D. A. (2021). A matter of taste: Gustatory sensitivity predicts political ideology. *Journal of Personality and Social Psychology, 121*(2), 394–409. https://doi.org/10.1037/pspp0000365

Shimizu, S., Hoyer, P. O., Hyvärinen, A., & Kerminen, A. (2006). A linear non-Gaussian acyclic model for causal discovery. *Journal of Machine Learning Research, 7*(72), 2003–2030. https://jmlr.org/papers/v7/shimizu06a.html

Shimizu, S., Inazumi, T., Sogawa, Y., Hyvärinen, A., Kawahara, Y., Washio, T., Hoyer, P. O., & Bollen, K. (2011). DirectLiNGAM: A direct method for learning a linear non-Gaussian structural equation model. *Journal of Machine Learning Research, 12*, 1225–1248. https://dl.acm.org/doi/10.5555/1953048.2021040

Shuman, E., Saguy, T., van Zomeren, M., & Halperin, E. (2021). Disrupting the system constructively: Testing the effectiveness of nonnormative nonviolent collective action. *Journal of Personality and Social Psychology, 121*(4), 819–841. https://doi.org/10.1037/pspi0000333

Spirtes, P., Glymour, C. N., & Scheines, R. (2000). *Causation, prediction, and search.* MIT Press.

Stelzl, I. (1986). Changing a causal hypothesis without changing the fit: Some rules for generating equivalent path models. *Multivariate Behavioral Research, 21*(3), 309–331. https://doi.org/10.1207/sl5327906mbr2103_3

Thoemmes, F. (2015). Reversing arrows in mediation models does not distinguish plausible models. *Basic and Applied Social Psychology, 37*(4), 226–234. https://doi.org/10.1080/01973533.2015.1049351

Thoemmes, F., & Lemmer, G. (2019). Mediation analysis revisited again. *Australasian Marketing Journal, 27*(1), 52–56. https://doi.org/10.1016/j.ausmj.2018.10.011

Methodology (METH) is the official journal of the European Association of Methodology (EAM).

PsychOpen GOLD is a publishing service provided by the Leibniz Institute for Psychology (ZPID), Germany.