Rethinking Statistics and Causality: Why Mechanisms Cannot Be Inferred from Projected Data Distributions

Egil Diau
Department of Computer Science
National Taiwan University
Taiwan, Taipei
egil158@gmail.com

Abstract

Statistical and causal inference have become universal currencies of explanation across the sciences, especially where underlying mechanisms remain opaque. Their authority rests on the assumption that patterns in observed data can reveal the processes that generated them. Yet persistent mismatches between empirical findings and real-world behavior point to a deeper limitation: observed data are projections of underlying systems, not the systems themselves. Such projections need not preserve the structural or semantic properties of what they represent. As a result, operations on projected data cannot be assumed to correspond to operations on the original structure. Statistical and causal inference often deepen this substitution by treating mathematical decomposition in the observed space as mechanistic decomposition of the system that produced it. But decompositions of projected data remain confined to the projected representation and are generally non-unique; they do not establish correspondence with the underlying mechanism. This reframes a central limit of modern inference: precision, fit, and decomposition within observed data are not evidence of mechanistic correspondence with the original structure. Mechanistic understanding therefore requires either direct operation on the underlying structure, or operation through a representation whose mapping has been shown to preserve the relevant properties of the original system, such as a validated simulation.

1 Introduction

Across many scientific domains, statistical and causal inference have become default tools for turning complex observations into interpretable results. Psychology Collaboration [2015], Tyner et al. [2026], economics Camerer et al. [2016], and biomedicine Ioannidis [2005], in particular, often rely on these methods when the underlying mechanisms are unknown or underspecified. Over time, this dependence has produced a sense of methodological certainty: if an analysis is statistically significant, or if a causal model is mathematically specified, the resulting claim appears to have explanatory force. Yet the growing number of failed replications, from behavioral studies Tyner et al. [2026] to clinical trials Begley and Ellis [2012], suggests that many such results may capture mathematical patterns in observed data rather than stable principles of the systems being studied.

The central problem is not insufficient data, weak statistical power, or poor experimental practice. It lies in the status of observed data itself. Observed variables are not the original structure of the system. They are projections of that structure into a measurable form. Unless the mapping from the system to the data is shown to preserve the relevant structure, operations on observed variables cannot be treated as operations on the system that generated them.

Figure 1: **Data as lower-dimensional projections of structure.** A single underlying system can give rise to multiple observational projections, each encoding a different low-dimensional representation of the same structure. Once projected and compressed, the relevant properties of the original system need not be preserved. Therefore, the underlying mechanisms are not identifiable from the projected data distribution alone; no operation on such distributions can by itself determine the original properties of the system.

This distinction is obscured because statistical and causal procedures operate symbolically on observed variables and their distributions. Variables can be correlated, regressed, adjusted, or arranged into graphical structures; distributions can be decomposed, conditioned, factorized, or fitted. But these operations remain operations on projected data. They do not by themselves show that corresponding operations exist in the underlying system that generated the observations.

This gives rise to a broader interpretive error: mathematical decomposition in projected data is treated as mechanistic decomposition in the original system. Mathematical decomposition is generally non-unique: the same observed relations can be represented through different bases, parameterizations, latent variables, or conditional factorizations. For this reason, even vague constructs or unrelated variables can yield a mathematical decomposition. Mathematical decomposition is therefore an algebraic arrangement, not evidence that the original system is organized in the same way.

We argue that mechanistic and causal understanding can follow only two routes. The first is to operate directly on the original structure itself. The second is to operate through a representation whose mapping preserves the relevant structure of the system being studied. Only under these conditions can statistical or causal analysis test properties of the original structure, rather than merely manipulate symbolic relations within projected data.

Our contributions.

- We argue that the replication crisis is a structural consequence of geometry: when projection does not preserve the original properties, true underlying relations can yield misleading statistical patterns, and spurious relations can produce apparently valid ones.
- We show that statistical and causal inference often treat mathematical decomposition in projected data as mechanistic decomposition of the original system. Yet such decomposition is performed within the projected data space, and mathematical decompositions are generally non-unique. The presence of a mathematical decomposition therefore does not establish correspondence with the mechanistic structure of the original system.
- We re-examine the foundations of statistical and causal inference, showing that correlation, significance, regression, and conditional dependence arose as geometric or algebraic operations on observed data but were later interpreted as explanations of mechanism—an interpretation their mathematics cannot support.

- We distinguish two routes for mechanistic and causal understanding: direct operation on the original structure itself, and operation through a structure-preserving representation. Both routes differ from symbolic manipulation of projected data, because they allow the relevant properties of the original structure to be tested.

2 Prerequisite: Data Collection, Projection, and the Substitution Error

2.1 Data Collection as Projection

Data collection is a form of projection. It records a richer underlying system through variables, measurements, categories, responses, or features. What enters the dataset is not the system itself, but a lower-dimensional or otherwise compressed representation of it.

This applies across ordinary empirical settings. An image is not the physical scene itself, but a two-dimensional record shaped by light, viewpoint, sensor properties, and preprocessing choices. A questionnaire response is not the psychological state itself, but a response constrained by wording, scale, and prompt. Market data are not the economic system itself, but recorded traces of transactions, prices, volumes, and institutional reporting rules. In each case, the dataset is a projected representation of a richer underlying system, not the system itself.

2.2 Mathematical Properties of Projection

Projection has several mathematical properties that are central to the argument:

- **Non-invertibility.** A projection generally does not admit a unique inverse. Once a richer underlying state is mapped into a lower-dimensional or compressed representation, the original state cannot be uniquely reconstructed from the observed representation alone.
- **Many-to-one mapping.** Projection can map multiple distinct underlying states to the same observed representation. The same observation may therefore be compatible with different underlying states.
- **Non-injectivity.** Formally, projection is generally non-injective: if $f(s_1) = f(s_2)$, it does not follow that $s_1 = s_2$. Different underlying states can become indistinguishable after projection.
- **Information loss.** Projection discards degrees of freedom. Components of the underlying system that are not preserved by the mapping are unavailable in the observed data, regardless of how precisely the observed data are later modeled.

These properties matter because empirical data collection is not a mathematical projection constructed to preserve the properties of the original system. In real-world settings, observations are produced through measurement devices, response formats, institutional records, coding schemes, and preprocessing choices. Such processes do not guarantee that the structural or mechanistic properties of the underlying system are carried into the observed representation.

2.3 The Substitution Error in Modern Inference

The substitution error begins when projected data are treated as the structure of the underlying system. Observed variables and distributions are then used as the basis for claims about mechanisms, causes, or structural relations. But inference is being carried out on the observed representation, not on the system that produced it.

The error deepens when projected observations are forced into an assumed formal structure, such as a linear model, a system of equations, a graphical model, or a directed acyclic graph. This creates two gaps. First, the assumptions are placed on projected observations rather than on the original system. Second, the assumed structure need not correspond to any organization in reality. The resulting quantities are therefore properties of the projected representation under those assumptions, not properties of the underlying system itself.

3 The Root of the Error: Inference from Projected Observations

In many empirical fields, modern inference has been shaped by three recurring pressures:

- **The underlying mechanisms were not specified.** Many fields operated without explicit accounts of how observations are produced. The processes responsible for generating these phenomena remained implicit, ill-defined, or absent from the modeling framework.
- **Conclusions were drawn from minimal observations.** Researchers routinely attempted to make claims about complex systems using small, sparse, or convenience datasets. Limited data were treated as sufficient to characterize the structure of the world.
- **The demand to extract results from vague constructs and weak effects.** Many fields began from broad constructs whose boundaries were unclear and effects whose magnitudes were weak, unstable, or difficult to interpret. Despite this, such constructs could still be operationalized, tested for association, and reported as empirical findings. In other words, obtaining a result did not require the construct to be clearly defined or the effect to be strong.

These pressures encouraged a substitution: data were treated not merely as evidence about the world, but as stand-ins for the processes that produced the observations. Samples were assumed to come from an underlying distribution; the distribution was assumed to reflect stable properties of the world; and stability was taken to imply mechanism. Once this substitution was accepted, inference inverted the logic of scientific explanation. Rather than specifying the underlying mechanisms, researchers attempted to infer them directly from projected observations.

3.1 The Limits of Statistical Inference

Statistical inference operates on projected observations rather than on the underlying system itself. Formally, what is recorded is not the system state $s \in \mathcal{S}$ but a compressed representation

$$x = f(s), \qquad x \in \mathbb{R}^n,$$

where f need not preserve the structural or semantic content relevant to the underlying mechanism. The distribution $P(x)$ therefore captures only the structure induced by this projection, not the mechanisms that generated the underlying states.

This distinction has a direct consequence for inference. Patterns in x—whether summaries, estimates, fitted distributions, or decompositions—remain confined to the projected representation. They do not reach $\mathcal{S}$, and their formal manipulation does not show correspondence with the structure that generated the observations. Increasing precision only sharpens operations within the observed space; it does not turn those operations into understanding of the underlying structure.

3.2 Algebraic Decomposition Is Not Mechanistic Decomposition

Once projected data are treated as if they were the original structure, a second error follows. Statistical and causal inference often decompose the observed representation mathematically and then treat this decomposition as evidence about the relationships or mechanisms in the world. But decomposing projected data is not the same as decomposing the structure that produced the data.

This error has two parts:

1. **Mathematical decomposition is not mechanistic decomposition.**
   A decomposition only rewrites an object in terms of a chosen basis, factorization, coordinate system, or set of variables. It does not show that the real system is organized in that way. Even weak, vague, or unrelated constructs can be decomposed mathematically, and because the basis, coordinate system, or parameterization is not fixed, the same observed distribution can admit arbitrarily many mathematical decompositions. Algebraic decomposability is therefore not evidence of mechanistic separability.
2. **The operations still occur within the projected data space.**
   If the projection from the underlying system to the observed data does not preserve the relevant properties of the original structure, then operations on the projected representation remain mathematical manipulations of that representation. They do not become operations on the underlying mechanism itself.

This applies regardless of the specific decomposition used. Regression, conditioning, residualization, factorization, dimensional reduction, and Bayesian factorization all operate on the observed representation. They rewrite, partition, summarize, or rearrange projected data into a particular form. But these operations do not show that the world itself contains the corresponding components, causes, or independent relations.

4 Statistical Inference as Geometry in Projected Data

Statistical inference attempts to derive relations in the original structure from manipulations of projected data. Since the underlying system is not directly available, traditional statistical methods translate observed relations into geometric relations within the data space. These relations take the form of angles, distances, projections, decompositions, and fitted patterns.

The deeper problem is that mathematical relations are not mechanistic relations. Yet in statistical practice, these relations are often treated as mechanisms, effects, or causal structures that the mathematics itself does not provide. This section examines this substitution through three representative statistical operations: correlation, significance, and regression.

4.1 The Geometry of Correlation

Pearson correlation Pearson [1896] is usually interpreted as a measure of how two observed variables vary together. In practice, it is often treated as if it captured a meaningful relation between the quantities being measured. Geometrically, however, correlation is simply the cosine of the angle between two centered data vectors in the observed data space:

$$r_{xy} = \frac{\mathrm{cov}(x, y)}{\sigma_x \sigma_y} = \frac{(x - \bar{x}) \cdot (y - \bar{y})}{\|x - \bar{x}\| \, \|y - \bar{y}\|} = \cos(\theta_{xy}).$$

Table 1: **Approximate angular separation for common correlation values.**

r_{xy}	**Cosine interpretation**	**Angle θ_{xy} (°)**
1.0	Perfect alignment	0
0.9	Very strong	26
0.8	Strong	37
0.7	Moderate–strong	46
0.5	Moderate	60
0.3	Weak	73
0.2	Very weak	79
0.0	None	90

As shown in Table 1, even moderate correlation coefficients correspond to surprisingly large angular separations. In psychology, typical reported correlations are around $r \approx 0.2$, which corresponds to an angular separation of roughly 79° Richard et al. [2003], Gignac and Szodorai [2016]. Recent large-scale replication evidence across the behavioral and social sciences shows an even sharper pattern: reported effects shrink from about $r = 0.25$ to $r = 0.10$, corresponding to a shift from roughly 75° to 84° Tyner et al. [2026]. In geometric terms, these variables are already nearly orthogonal. Yet such relations are still routinely treated as meaningful effects rather than weak angular alignments inside a projected data space.

The deeper issue is that an angle between quantities from different representational spaces has no well-defined meaning. In empirical research, variables often originate from different domains, measurement procedures, or representational spaces. Once placed into the same data table, correlation assigns them a geometric relation, but that relation does not by itself specify any mechanistic connection. A familiar example is the reported correlation between chocolate consumption and Nobel Prize counts: two quantities with no clear mechanistic relation can still appear geometrically aligned Messerli [2012]. The problem therefore extends beyond weak correlations: measuring an angle between variables from different representational spaces does not provide a mechanistic interpretation of the structure that produced them.

4.2 Significance as One-Dimensional Evidence

Significance testing is usually treated as a rule for deciding whether an observed pattern is unlikely under a null model Fisher [1970], Neyman and Pearson [1933]. Formally, the p-value is a tail probability in the space of a summary statistic:

$$p = P(T(X) \geq t_{\text{obs}} \mid H_0).$$

This quantity lives within the space defined by the statistic, the null model, and the reference distribution. It therefore reflects a position inside a statistical construction, not the structure of the system that produced the data.

The substitution occurs when this one-dimensional position is treated as evidence about a high-dimensional system. Thresholding converts location within a reference distribution into a semantic claim—"significant," "real," or "meaningful." But the mathematics supplies only a relation between a statistic and a null model. It does not supply correspondence with the real structure of the data, and it does not supply the semantic meaning later assigned to the threshold.

In classical settings, this calibration is often justified through asymptotic convergence. Under i.i.d. observations with finite variance, the central limit theorem gives

$$\sqrt{n}\big(T(X) - \theta\big) \xrightarrow{d} \mathcal{N}(0, \Sigma),$$

This is a mathematical guarantee under specified assumptions, not an empirical guarantee that the reference distribution corresponds to the system being studied. When those assumptions do not map onto the real structure of the data, the resulting language of significance has no established correspondence with the empirical world.

4.3 Regression as Projection

Regression is usually interpreted as estimating the effect of one variable on another. Geometrically, however, the fitted outcome is the projection of the outcome vector onto the subspace spanned by the predictors Rencher and Schaalje [2008], Hastie [2009]. In the simplest case, the regression coefficient is given by

$$\beta = \frac{\operatorname{cov}(x, y)}{\operatorname{var}(x)} = \frac{(x - \bar{x}) \cdot (y - \bar{y})}{\|x - \bar{x}\|^2}.$$

This coefficient describes the coordinate of y's projection along the direction of x under a least-squares criterion. It is a mathematical relation between observed vectors, not an effect supplied by the formula itself.

The substitution occurs when this fitted relation is treated as the causal effect of x on y. Nothing in the coefficient specifies influence, direction, or mechanism. Different underlying systems can produce the same regression coefficient, and the same system can yield different coefficients under different choices of variables, controls, measurements, or representations. The coefficient therefore belongs to the geometry of the fitted data space, not to the original structure of the system.

5 Causal Inference as Algebra on Projected Data

Causal inference seeks to identify relations of intervention and effect in the original structure of a system Pearl [2009], Hernán and Robins [2010], Peters et al. [2017]. In practice, however, it operates on observed variables and their distributions, which are already projected representations of that structure. Conditioning, adjustment, factorization, counterfactual comparison, graphical separation, and intervention calculus are therefore algebraic operations within the projected data space.

The central problem is that causality belongs to the original structure, not to its projection. It concerns how the underlying system changes when part of that system is changed. Unless the projection preserves the relevant structure, operations on observed variables remain operations on the projected representation, regardless of how sophisticated the causal method appears.

5.1 Probability Operations Do Not Ensure Real-World Correspondence

A rarely examined gap lies at the foundation of probability operations. Multiplication and division are often treated as ordinary ways to combine information about the world. But the notation hides a crucial distinction: a probability expression can be written within a formal system without showing that the corresponding relation exists in the world.

Multiplication already shows the problem. In probability theory, two probabilities can be multiplied only when the two probabilities are treated as independent:

$$P(A)P(B) = P(A \cap B).$$

This condition allows the probability product to be identified with the probability of a set-level operation, namely the intersection $A \cap B$. But independence itself is not a set-level operation, nor is it a property of the world. It is a formal condition that makes calculation tractable by allowing a product of probabilities to stand in for the probability of an intersection. Therefore, in real-world settings, the product of two probabilities does not necessarily correspond to the probability of their intersection.

Division creates a deeper problem. Under independence, probability multiplication can at least be identified with the probability of the intersection $A \cap B$. Probability division has no analogous set-level operation. For example, $P(A)/P(B)$ does not correspond to any event obtained by "dividing" A by B. In practice, it is unclear what real-world operation this division is supposed to represent.

Conditioning and Bayes' theorem inherit this problem. Conditional probability is defined through division Kolmogorov [2018],

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)},$$

and Bayes' theorem follows by rearranging the same definition de Laplace [1820]:

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}.$$

The problem is that probability division has no clear real-world counterpart: it is unclear what operation in the world corresponds to dividing one probability expression by another. Meanings such as learning, belief revision, information updating, or causal transformation are then attached afterward. These meanings are interpretive additions, not consequences supplied by the algebra itself.

The same problem extends to conditional factorization. A joint distribution may be rewritten as a product of conditional terms in graphical models Pearl [2014], Lauritzen [1996], Koller and Friedman [2009],

$$P(X_1, \ldots, X_n) = \prod_i P(X_i \mid \mathrm{Parents}(X_i)).$$

But each conditional term still relies on probability division. Factorizing a distribution therefore does not escape the earlier problem; it only arranges the same unclear operation into a structured expression. Such a factorization does not show that the variables are causal parents, separable mechanisms, or real transformations in the world. Conditional factorization therefore should not be mistaken for causal identification.

5.2 Causal Models Still Operate on Projected Data

Causal inference methods differ in notation, assumptions, and implementation, but they share a common dependence on observed variables and their distributions. Structural equation modeling Bollen [2014], graphical causal models Pearl [2009], intervention calculus Pearl [2009], and potential-outcome frameworks Rubin [1974] construct causal claims through formal operations on these objects. The difficulty is that observed variables are already projected representations of an underlying system. Operations defined over them therefore do not, by themselves, reach the structure that generated the observations.

This limitation becomes decisive once projection is made explicit. Projection compresses an underlying system into observed variables, and need not preserve the properties required for causal identification. If those properties are not preserved, then no manipulation within the projected data space can determine the causal relations of the original structure.

Machine-learning-based causal approaches Athey and Imbens [2016], Johansson et al. [2016], Chernozhukov et al. [2018] inherit the same limitation. Replacing explicit causal notation with learned representations or pattern-based procedures does not change the object on which the operation is performed: it remains projected data. If this representation does not preserve the relevant properties of the original structure, then causal relations inferred within it cannot be assumed to correspond to causal relations in the original system.

Causal identification therefore requires access to the original structure itself, or to a representation that preserves the relevant properties of that structure. Without such correspondence, causal modeling remains manipulation within projected space rather than identification of causality in the original system.

6 Mechanistic Understanding Requires Structure-Preserving Representation

The original structure of a real-world system lies beyond the observed data: it is a higher-dimensional system whose states and relations give rise to the observed phenomenon. Mechanism and causality lie in how this underlying system responds to changes or interventions. To understand them, the object being manipulated must preserve the relevant structure of the original system; otherwise, the operation remains confined to a projection whose properties may differ from those of the system it represents.

Mechanistic understanding therefore can be grounded in two ways. It can come from direct operation on the underlying structure itself, or from operation through a representation that preserves the relevant properties of that structure.

- **Direct operation on the underlying structure.**
  The first route is to intervene on, measure, or observe the underlying system itself. This approach aims to characterize the structure and mechanisms that give rise to the observed phenomenon. Controlled interventions and physical measurement belong to this route: they seek to show how changes in the underlying structure produce changes in the observed phenomenon.
- **Operation through a structure-preserving representation.**
  The second route is to operate through a representation whose mapping preserves the relevant properties of the original system. Such a representation becomes mechanistically meaningful only when its internal operations remain tied to the structure of the system being explained. A validated simulation is one concrete way to satisfy this condition: it can support mechanistic understanding when manipulations inside the simulation correspond to changes in the underlying system.

The central distinction is therefore whether a representation preserves the relevant structure of the original system. If it does, operations within the representation can support mechanistic explanation. If it does not, the model remains a manipulation of projected data, regardless of how expressive or accurate it appears.

7 Discussion: Machine Learning and the Original Structure

The issue is not data-driven modeling itself, but what such modeling is taken to establish. Machine learning, including deep learning, can be used for prediction or approximation within the observed data space without claiming to identify the mechanisms, causes, or structural properties of the underlying system. It learns regularities in the projected representation.

The problem arises when statistical or causal inference is used to infer properties of the original structure from projected observations. Once inference is carried out within the projected data space, its operations are defined over the observed representation rather than the underlying system itself. They therefore do not by themselves establish the corresponding mechanisms, causal relations, or semantic structure of the original system.

8 Conclusion

The limitation identified here is structural: mechanisms and causal relations belong to the original structure of a system, not to its projected representation. Observed data are projections of underlying systems, and such projections need not preserve the structure of the systems they represent. As a result, operations performed within the observed space—such as fitting, decomposition, and probabilistic manipulation—remain confined to the projected representation. They do not operate on the original structure, nor do they establish correspondence with it.

This limitation is not resolved by increasing data, improving estimation, or refining formal models, because it arises from the mapping between the original system and its observed representation. When that mapping does not preserve the relevant structure, no operation within the projected space can determine relations in the original structure.

Mechanistic understanding therefore requires either direct operation on the underlying structure, or operation through a representation whose mapping to that structure has been independently shown to preserve the relevant properties.

Declaration of LLM Usage

The authors used OpenAI's ChatGPT to assist in refining phrasing and improving clarity. All theoretical arguments and interpretations are original and authored by the researchers.

References

S. Athey and G. Imbens. Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences*, 113(27):7353–7360, 2016.

C. G. Begley and L. M. Ellis. Raise standards for preclinical cancer research. *Nature*, 483(7391): 531–533, 2012.

K. A. Bollen. *Structural equations with latent variables*. John Wiley & Sons, 2014.

C. F. Camerer, A. Dreber, E. Forsell, T.-H. Ho, J. Huber, M. Johannesson, M. Kirchler, J. Almenberg, A. Altmejd, T. Chan, et al. Evaluating replicability of laboratory experiments in economics. *Science*, 351(6280):1433–1436, 2016.

V. Chernozhukov, D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, W. Newey, and J. Robins. Double/debiased machine learning for treatment and structural parameters, 2018.

O. S. Collaboration. Estimating the reproducibility of psychological science. *Science*, 349(6251): aac4716, 2015.

P. S. de Laplace. *Théorie analytique des probabilités*, volume 7. Courcier, 1820.

R. A. Fisher. Statistical methods for research workers. In *Breakthroughs in statistics: Methodology and distribution*, pages 66–70. Springer, 1970.

G. E. Gignac and E. T. Szodorai. Effect size guidelines for individual differences researchers. *Personality and individual differences*, 102:74–78, 2016.

T. Hastie. The elements of statistical learning: data mining, inference, and prediction, 2009.

M. A. Hernán and J. M. Robins. *Causal inference*. CRC Boca Raton, FL, 2010.

J. P. Ioannidis. Why most published research findings are false. *PLoS medicine*, 2(8):e124, 2005.

F. Johansson, U. Shalit, and D. Sontag. Learning representations for counterfactual inference. In *International conference on machine learning*, pages 3020–3029. PMLR, 2016.

D. Koller and N. Friedman. *Probabilistic graphical models: principles and techniques*. MIT press, 2009.

A. N. Kolmogorov. *Foundations of the theory of probability: Second English Edition*. Courier Dover Publications, 2018.

S. L. Lauritzen. *Graphical models*, volume 17. Clarendon press, 1996.

F. H. Messerli. Chocolate consumption, cognitive function, and nobel laureates, 2012.

J. Neyman and E. S. Pearson. Ix. on the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 231(694-706):289–337, 1933.

J. Pearl. *Causality*. Cambridge university press, 2009.

J. Pearl. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Elsevier, 2014.

K. Pearson. Vii. mathematical contributions to the theory of evolution.—iii. regression, heredity, and panmixia. *Philosophical Transactions of the Royal Society of London. Series A, containing papers of a mathematical or physical character*, (187):253–318, 1896.

J. Peters, D. Janzing, and B. Scholkopf. *Elements of causal inference: foundations and learning algorithms*. MIT press, 2017.

A. C. Rencher and G. B. Schaalje. *Linear models in statistics*. John Wiley & Sons, 2008.

F. D. Richard, C. F. Bond Jr, and J. J. Stokes-Zoota. One hundred years of social psychology quantitatively described. *Review of general psychology*, 7(4):331–363, 2003.

D. B. Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, 66(5):688, 1974.

A. H. Tyner, A. L. Abatayo, M. Daley, S. Field, N. Fox, N. A. Haber, K. M. Hahn, M. K. Struhl, B. Mawhinney, O. Miske, et al. Investigating the replicability of the social and behavioural sciences. *Nature*, 652(8108):143–150, 2026.