Big Data and Causation – Reply to Serena Galli

Wolfgang Pietsch (wolfgang.pietsch@tum.de)

Accounts of induction based on a difference making logic have a long history in debates on scientific method (e.g. Bacon 1620, Bk. II; Herschel 1830, P. II Ch. VI; Mill 1843, Bk. III Ch. VIII; Mackie 1967, Appendix; Baumgartner & Graßhoff 2004). This type of induction has been called variational or variative induction (Cohen 1989; cf. Russo 2007, 2009). I respond to objections raised by Serena Galli (2023) against the specific account of variational induction elaborated in Pietsch (2016; 2022, Chs. 5,6; a formal summary is provided in 2026a) and its application to the analysis of big data and machine learning practices (Pietsch 2021; 2026b). These objections primarily concern the directedness of causation and situations of causal underdetermination. While these present genuine, non-trivial problems, I argue that variational induction can address them, provided one accepts certain—admittedly controversial—assumptions.

1. Causal underdetermination problems

1.a) Overdetermination and pre-emption

Galli argues that variational induction is affected by underdetermination. An underdetermination problem as understood by Galli arises when an inferential framework cannot distinguish between different causal models. According to a first underdetermination problem raised in Section 1 i) of Galli (2023), variational induction is not capable of distinguishing a situation of causal pre-emption from a situation of causal overdetermination. Also, variational induction allegedly fails to account for the asymmetry between pre-empting and pre-empted cause.

The short answer to this underdetermination problem is that, according to variational induction, different causal models are indeed compatible with any given evidence. However, these causal models all agree with regards to difference making and therefore in their assessment of manipulability and predictability. In other words, there is *no* underdetermination in terms of difference making with respect to a given background and thus regarding manipulability and predictability of variables with respect to this background.

Thus, the reply to the first underdetermination problem draws on the following two premises:

1) *Radical context-dependence*: any causal statement or model must always be related to a homogeneous background or context B. In other words, there are no context-independent causal statements. Accordingly, assessments of causal relevance and irrelevance may change when different backgrounds are considered.

2) *Instrumentalism about causation*: causal underdetermination should ultimately be judged in terms of whether there is underdetermination with respect to difference making. Importantly, if different causal models imply the same difference-making statements, these models will also agree concerning their assessment of manipulability and predictability.

In the causal underdetermination problem raised in Section 1 i) of Galli (2023), the evidence consists of the following four states sharing the background B: i) A1 ∧ A2 ∧ C ∧ B, ii) A1 ∧ ¬A2 ∧ C ∧ B, iii) ¬A1 ∧ A2 ∧ C ∧ B, iv) ¬A1 ∧ ¬A2 ∧ ¬C ∧ B. Underdetermination results, because variational induction cannot determine based on this evidence whether in situation i) both A1 and A2 are overdetermining "causes" of C or whether e.g. A2 is pre-empted by A1 and thus only A1 is a "cause" of C in situation i). The causal model is underdetermined even though the given evidence concerns all possible combinations of the considered circumstances A1 and A2.

However, there is *no* underdetermination in terms of difference making with respect to the given background B. In particular, A1 is a difference maker to C in the absence of A2 and A2 is a difference maker to C in the absence of A1, while neither A1 nor A2 are difference makers to C in the presence of the other circumstance. With respect to background B, these difference-making statements hold independently of whether the "underlying" causal structure is overdetermination or pre-emption. Since, according to variational induction, difference making alone determines any assessment of manipulability and predictability of variables, there is *no* underdetermination regarding manipulability and predictability either. Thus, drawing on premises 1) and 2) above, there is no need to distinguish pre-emption from overdetermination with respect to background B if one is only interested in an assessment of manipulability and predictability of these variables based on each other.

As Galli correctly points out, considering further variables, e.g. mediating variables, could help distinguishing between pre-emption and overdetermination, but may itself lead to still further underdetermination situations in terms of different causal models compatible with the given evidence. However, also these further underdetermination situations do *not* imply underdetermination regarding difference making and thus all compatible causal models will always agree on their assessment concerning manipulability and predictability with respect to the given background.

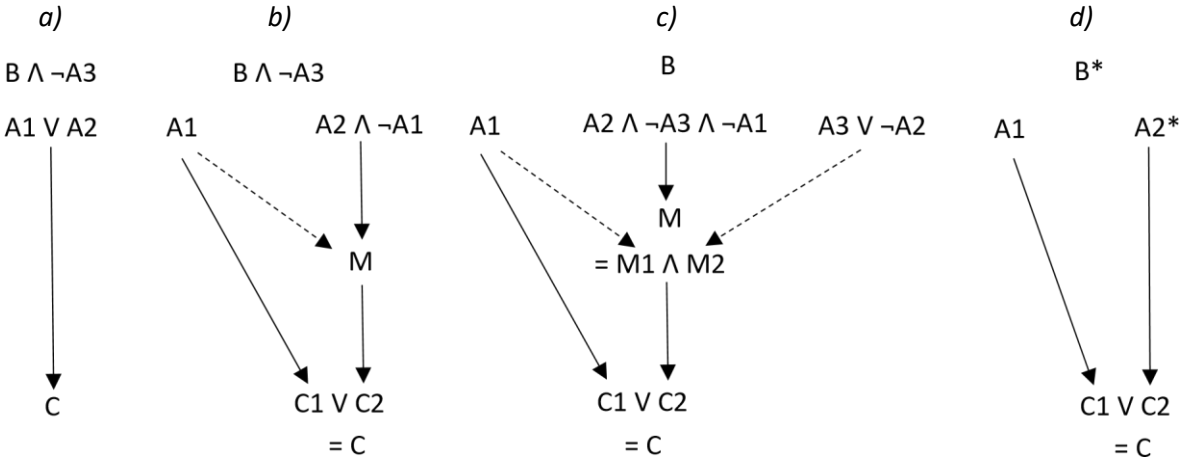

Figure I: a) underdetermination regarding pre-emption and causal overdetermination; b) preemption of A2 by A1; c) pre-emption with external manipulability of mediating variable M; d) overdetermination

The first underdetermination problem is illustrated in Figure I, where I am using a slightly different representation than Galli (2023). Non-dashed arrows denote causal relevance between variables, i.e. essentially a bi-conditional statement between two variables A and B: "if and only if (iff) A, then B". Dashed arrows denote causal relevance between the variable at the base of the arrow and the *negation* of the variable at the tip of the arrow. "Λ" denotes logical „and", "V" denotes logical „or", and "¬" denotes negation.

Figure I.a) shows the underdetermination situation where the causal relevance of alternative causes A1 V A2 to phenomenon C is compatible both with an underlying model of pre-emption, in which one of the alternative causes A1 or A2 pre-empts the other, and with a model of overdetermination, in which A1 and A2 imply C independently of each other. More detailed models of pre-emption are shown in Figures I.b) and I.c), and of causal overdetermination in Figure I.d).

To illustrate the proposed solution to the underdetermination problem, let me discuss a well-known example of pre-emption: a desert traveller has two enemies, one poisons the water of the traveller, the other drills a hole in the water bottle of the traveller. The traveller will eventually die of thirst

having no drinking water due to the hole in his water bottle. But without the hole, the desert traveller would have drunk poisoned water, which would have killed him anyway.

In Figure I, A1 denotes the event "making a hole in the water bottle", A2 the event "poisoning the water in the water bottle", A3 the event "putting an antidote in the water bottle, which neutralizes the poison", M the mediating event "drinking poisoned water out of the bottle" with M1 denoting the event "drinking water out of the bottle" and M2 the condition "water is poisoned at the moment of drinking", and C the event "death of the traveller" with C1 being "dying of thirst" and C2 being "dying of poisoning".

As mentioned, non-dashed arrows denote causal relevance between the corresponding events. For example, in Figure I.b), M is causally relevant to C2 (iff the traveller drinks poisoned water, he will die of poisoning) and A1 is causally relevant to C1 (iff there is a hole in the bottle, the traveller will die of thirst). Dashed arrows denote causal relevance between the event at the base of the arrow and the negation of the event at the tip of the arrow. For example, in Figure I.c), A1 is causally relevant to not-M1 (iff there is a hole in the bottle, the traveller can't drink water out of the bottle) and A3 V ¬A2 is causally relevant to not-M2 (iff there is an antidote in the water or if the water has not been poisoned in the first place, the water is not poisoned at the moment of drinking).

The causal-relevance relations of Figures I.b) and I.c) are incompatible with the overdetermination situations of both Figures 1b and 1e of Galli (2023). Importantly, Figures I.b) and I.c) illustrate the way a practitioner would proceed in order to distinguish a situation of pre-emption from overdetermination, namely by first identifying a mediating variable, here the drinking of the poisoned water (Figure I.b), and second by finding a way to manipulate the mediating variable independently of the pre-empting cause (Figure I.c). In this manner, M can be identified as a mediating variable between A2 and C, thereby excluding causal model 1e.

Note that both Figures I.b) and I.c) show asymmetries between pre-empting and pre-empted cause. For example, the pre-empting cause A1 is causally relevant to C1 independently of A2, while the preempted cause A2 requires the absence of A1 for causal relevance to C2. Figures I.b) and I.c) prove that whether there is symmetry or not crucially depends on the chosen representation, in particular on the background and on which variables are taken into account.

By contrast, situations of overdetermination can be identified, when a mediating variable of the type illustrated in Figures I.b) and I.c) cannot be found. In Figure I.d), such an overdetermination situation is shown for comparison. A2* could refer to a different kind of poisoning that is independent of A1, for example "poisoning by a poisoned syringe".

In summary, while there is underdetermination in terms of different causal models that are compatible with the evidence, there is *no* underdetermination in terms of difference making with respect to the given background and therefore *no* underdetermination in terms of manipulability and predictability between variables. Situations of pre-emption and of overdetermination can be distinguished by taking into account mediating variables that are independently manipulable. As emphasized, the bullet to be bitten is that *any* causal model must always be related to a specific background satisfying homogeneity (premise 1 above).

1.b) Epiphenomena

The second objection raised in Section 1 ii) of Galli (2023) concerns epiphenomena and essentially leads to an underdetermination problem quite similar to the one discussed above. For example, Figure 2 of Galli (2023) depicts various causal models compatible with the evidence of observations 9 and 10. Due to these similarities, the second objection can be approached in an analogous manner by stressing context-dependence, instrumentalism about causation and as a further premise:

3) *Strict determinism*: if there is a single exception, i.e. a single contravening observation, then there is no causal relevance or causal irrelevance with respect to the given background.

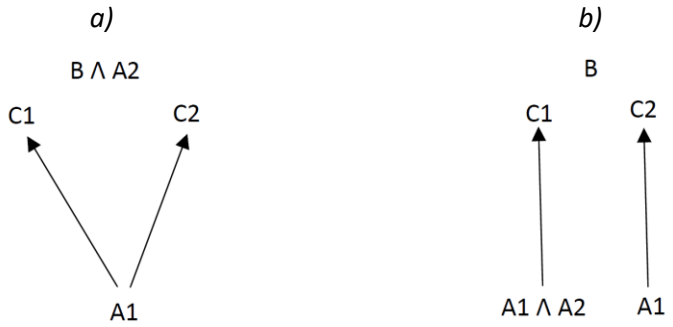

Figure II: a) C1 as epiphenomenon or proxy variable correlating with C2; b) external manipulation of the proxy variable by means of variable A2

The evidence underlying the causal model of Figure II.a) consists in the following two states sharing the background B ∧ A2: i) A1 ∧ C1 ∧ C2 ∧ B ∧ A2, ii) ¬A1 ∧ ¬C1 ∧ ¬C2 ∧ B ∧ A2. Underdetermination results, because, based on this evidence, variational induction cannot distinguish between various causal models with respect to background B ∧ A2, for example whether A1 is a "cause" of C1 and a "cause" of C2 or whether C1 is a mediating "cause" between A1 and C2, etc. Figure II.a) shows a causal model, according to which C1 is an epiphenomenon that correlates with C2, but is not a "cause" of C2.

However, also in this case, the underdetermination is not problematic since there is *no* underdetermination in terms of difference making and therefore *no* underdetermination in terms of manipulability and predictability. Indeed, according to the causal model depicted in Figure II.a), even though C1 is merely an epiphenomenon, the variable C2 can only be changed by a corresponding change of variable C1. This is because the causal-relevance relations of Figure II.a) imply biconditional relationships, i.e. iff-then relationships. Therefore, strictly speaking, manipulating C1 with respect to the background B ∧ A2 necessarily leads to a corresponding change of C2. While this may sound odd, it reflects the strict background dependence of variational induction. In fact, all causal models of Figures 2a to 2f of Galli (2023) lead to the same difference-making statements (cp. Pietsch 2022, Sect. 5.4.5).

The different situations of Figure 2 can be distinguished by considering further variables and by extending the background. For example, in Figure II.b), the background is extended from B ∧ A2 to B, i.e. to a background where A2 is allowed to change. Here, A2 is a variable, with which C1 can be externally manipulated. By taking into account A2, causal models can be excluded, in which C1 is a mediating variable between A1 and C2 with respect to background B. Also, causal models can be excluded in which C1 is causally relevant to C2 with respect to background B. However, C2 ∧ A2 now turns out to be causally relevant to C1 with respect to background B. Thus, as in the case of the first underdetermination problem and as Galli correctly emphasizes, certain ambiguities always remain. But again, none of these ambiguities implies an underdetermination with respect to difference making and thus with respect to manipulability or predictability.

Let us look at an example for illustration. Let A1 be the event "it rains", C1 the event "the street is wet", C2 the event "Anna wears her raincoat", and A2 the state of "Anna being outdoors". Now, an innocent observer from another planet not familiar with the weather and people's behaviour on Earth may at first think that Anna wearing her raincoat is a cause of the street being wet, if in all observations Anna is always outdoors. But then the observer notices that Anna never wears her raincoat when she is indoors and has to conclude that C1 in contrast to A1 is not causally relevant to C2.

This strategy broadly corresponds to how a practitioner would proceed. By trying to independently manipulate certain variables, specific causal models can be excluded. Epiphenomena can eventually be identified by distinguishing robust causal relationships holding with respect to a broad background from those that are easily disturbed by means of external variables. As highlighted by premise 3),

variational induction is ultimately a deterministic framework. Thus, a single counterexample, e.g. a single observation iii) ¬A1 Λ C1 Λ ¬C2 Λ B Λ A2 or iv) A1 Λ ¬C1 Λ C2 Λ B Λ A2 in comparison to a large number of observations i) A1 Λ C1 Λ C2 Λ B Λ A2 and ii) ¬A1 Λ ¬C1 Λ ¬C2 Λ B Λ A2, still suffices to rule out causal relevance of A1 or C2 to C1 with respect to background B Λ A2 since the requirement of homogeneity would not be fulfilled. Therefore, Figure 2g of Galli (2023) is not an adequate causal model for the given evidence i) and ii) as it implies the existence of such contravening observations iii) or iv).

1.c) Causal asymmetry

Galli correctly recognizes that the definition of causal relevance is largely symmetric with respect to antecedent A and consequent C—in particular because causal relevance describes a one-to-one correspondence between antecedent and consequent. Note that there is a potential asymmetry in the distinction of the terms “condition” or “circumstance” A and “phenomenon” C. Various problems regarding the direction of time and of causation have plagued physics and philosophy for ages and it is not plausible that a solution will be forthcoming. However, the situation is not as hopeless as Galli suggests. I will sketch in the following some preliminary ideas based on the notion of causal chains how to introduce causal asymmetry in the framework of variational induction (cp. Pietsch 2022, Sec. 5.2.3 based on Pietsch 2016, Sec. 2c). The simplest and at least partial solution, which is also mentioned by Galli, is to introduce time indices for all events in combination with the premise that effects are always posterior to their causes. Such a causal asymmetry could be implemented as an additional constraint, e.g. by means of respective definitions of the terms “conditions” and “phenomenon”.

There exists a related difficulty not explicitly mentioned in Galli (2023) which concerns *causal ordering*. In addition to the problem of *causal asymmetry*, in some evidence situations, variational induction is not capable of distinguishing different orderings of causally related events. For example, the evidence i) C1 Λ C2 Λ C3 Λ B Λ ¬I1 Λ ¬S2; ii) ¬C1 Λ ¬C2 Λ ¬C3 Λ B Λ ¬I1 Λ ¬S2 is compatible with C1 being a mediating event between C2 and C3, C2 being a mediating event between C1 and C3 or C3 being a mediating event between C1 and C2 with respect to background B Λ ¬I1 Λ ¬S2. In other words, the causal ordering of the events C1, C2, C3 cannot be determined by variational induction based on the mentioned evidence. So, we are again faced with an underdetermination problem.

However, as in the other cases discussed so far, the situation can be improved by taking into account further variables and by considering a more extensive background. In my opinion, the approach delineated below follows largely, how a practitioner would determine causal ordering and causal direction. Central concepts are interruption variables I, which erase causal links between variables, and stimulation variables S, which externally stimulate variables.

A first example depicted in Figure III concerns only causal ordering and mostly lacks a causal direction. C1, C2, and C3 could each represent the turning of a respective toothed wheel. A double arrow between variables denotes that causal relevance can hold in both directions between the variables. In the causal model shown in Figure III.a), toothed wheel C2 is interlocked with toothed wheel C1 and with toothed wheel C3, but there is no direct interaction between C1 and C3. In order to determine the causal ordering of C1, C2, and C3, interruption events can be employed, realized e.g. by moving the respective toothed wheels apart.

Figure III.b) shows the setup of Figure III.a) with interruption event I1 being active—as accounted for by the respective background. Interruption event I1 eliminates the causal relevance relation between C1 and C2, while leaving the causal relevance relation between C2 and C3 intact. Interruption event I2 eliminates the causal relevance relation between C2 and C3, while leaving the causal relevance relation between C1 and C2 intact. Examining these two different interruption events allows establishing the causal ordering between the events C1, C2, C3.

In Figure III.c), stimulation event S2 activates an additional variable A2 outside the causal chain C1 – C2 – C3 with A2 being causally relevant to C2.

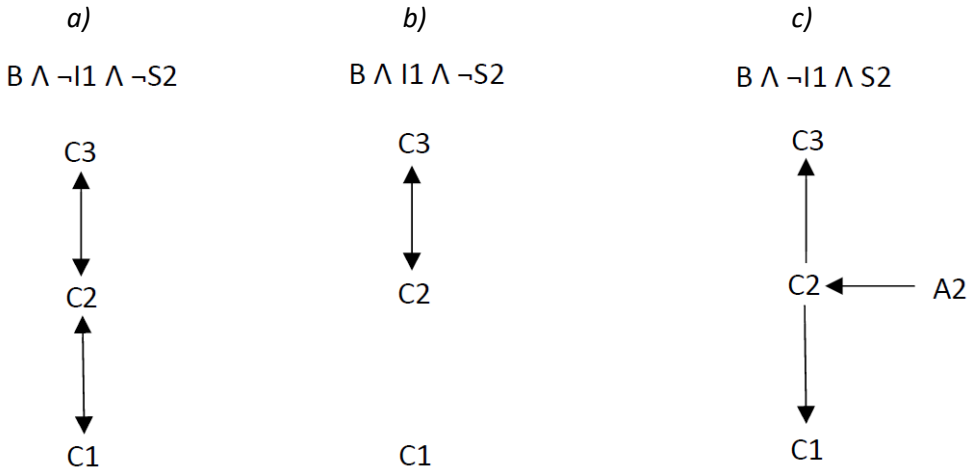

Fig. III showing a causal chain C1 - C2 - C3 of "toothed wheels": in a) C2 mediates between C1 and C3; in b) the interruption variable I1 eliminates the causal relevance relation between C1 and C2; in c) stimulation variable S2 activates an additional variable A2 outside the causal chain, with which variable C2 can be directly manipulated.

Figure III shows causal models with undirected, symmetric ("simultaneous") interaction, allowing for causal-relevance relations in both directions between variables, as in the toothed-wheels example. By contrast, Figure IV shows causal models of directed or asymmetric causal-relevance relations, which do not allow "backwards" causal relevance. For example, each event C1, C2 and C3 could represent a collision between billiard balls stopping the previous billiard ball and setting the respective billiard ball in motion towards a still further billiard ball. If the collisions are inelastic, the billiard balls all have the same weight and move in the same direction, only one billiard ball will be moving at any time.

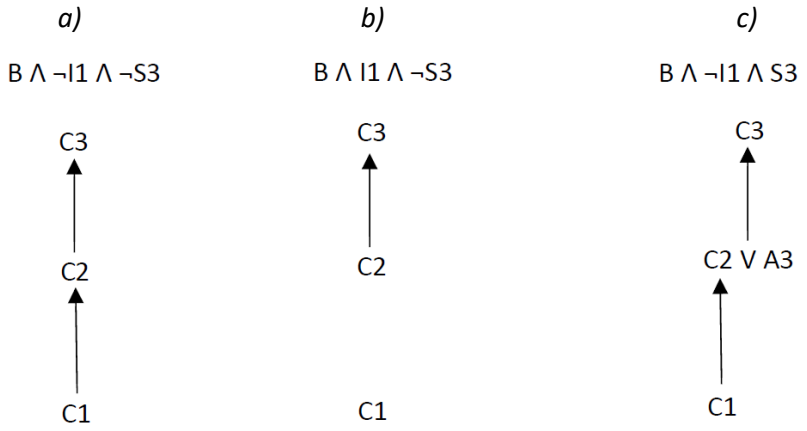

Fig. IV showing a causal chain C1 – C2 – C3 of "inelastic billiard ball collisions": in a) the causal relevance relations are directed due to the underlying time asymmetry; in b) causal ordering is determined based on interruption variable I1; in c) causal asymmetry is determined by means of stimulation variable S3 activating variable A3, which allows for external manipulation of C3. Manipulation of C3 by means of A3 leaves causally prior variables C2 and C1 unchanged.

Figure IV.a) shows a causal chain C1 – C2 – C3 with directed causal-relevance relations. However, based on the underlying evidence i) C1 ∧ C2 ∧ C3 ∧ B ∧ ¬I1 ∧ ¬S3; ii) ¬C1 ∧ ¬C2 ∧ ¬C3 ∧ B ∧ ¬I1 ∧ ¬S3, ordering and direction between the events C1, C2, and C3 are again underdetermined with respect to background B ∧ ¬I1 ∧ ¬S3.

As in the example of the toothed wheels, the ordering of these events can be determined by means of interruption events I1, I2 eliminating different causal relevance relations (see Figure IV.b). Interruption events might for example consist in putting a fixed obstacle in the path between the

respective billiard balls. However, such interruption events do not allow determining a causal asymmetry between C1, C2, and C3.

Stimulation variables S1, S2, and S3 activate additional variables A1, A2, and A3, which externally stimulate the respective events C1, C2, and C3, i.e. from outside the causal chain (see Figure IV.c). Such external stimulation allows establishing causal asymmetry. A stimulation event could be pushing the respective billiard ball with a cue stick. In contrast to the "toothed wheels", the stimulation events of the "billiard balls" example are *not* causally relevant to events causally prior in the chain. For example, if billiard ball C3 is stimulated by the cue stick A3, this will have no impact on billiard balls C2 and C1, but only on subsequent billiard balls C4, C5 etc. which are set in motion by C3. Note that stimulation events allow determining both causal ordering as well as causal asymmetry.

In summary, stimulation events allow establishing causal asymmetry because they do not affect events that lie "backwards" in the causal chain. Such asymmetric causal chains may serve as a kind of causal clock, which can be used to determine causal asymmetries in the relationships between other variables.

2. Big Data and Causal Knowledge

The fact that the analysis of big data by machine learning algorithms enables reliable predictions and effective interventions in the world supports the assumption that these algorithms at least sometimes correctly identify causal relationships. In addition, there is a more profound argument for the causal nature of big data models. As argued for example in Pietsch 2021, Sect. 4.1.3, variational induction relies on difference making under homogeneity. Accordingly, variational induction yields causal relationships that can be used for manipulating phenomena. As shown in Pietsch 2021, Sect. 4.2, some of the most successful machine learning algorithms dealing with big data also rely on such difference making under homogeneity. However, they typically implement difference making not in the pure form of variational induction, but in an approximate manner. Also, homogeneity may not always hold strictly. Therefore, depending on the prepared data set and the specific algorithm, these machine learning algorithms in general will not identify relations of causal relevance and irrelevance, but only approximations, e.g. the relationship between a good proxy and the phenomenon of interest. Therefore, the model extracted by a machine learning algorithm may serve only for prediction and might not be useful for effective intervention. However, this approximate implementation of variational induction by machine learning algorithms supports my claim that these algorithms yield models that approximate an underlying causal structure.

Galli criticizes that variational induction blurs several important conceptual distinctions, most notably between correlation and causation and between manipulability and predictability. I somewhat agree, but do not find this particularly problematic as it seems to fit rather well with scientific practice. For example, I argued above that with respect to a certain background, an epiphenomenon may turn out to be causally relevant to the phenomenon of interest, essentially because with respect to that background, the epiphenomenon can only be changed by changing a common cause variable that is also causally relevant to the phenomenon of interest. In other words, manipulating the alleged epiphenomenon necessarily leads to a change of the phenomenon of interest. Of course, this assessment will change, when different backgrounds are considered, which will eventually allow identifying the epiphenomenon as such. Thus, the distinction between manipulability and mere predictability based on an epiphenomenon is to some extent context-dependent. However, it seems to be a common experience that manipulation of certain variables is more or less effective for changing a phenomenon of interest, i.e. there is a large grey area instead of a clear distinction between variables that can be used for manipulation and others that cannot.

I maintain that causal knowledge is primarily a guide for effective intervention in the world. However, causal models can also establish correlations, which can be used for prediction, e.g. between the phenomenon of interest and an epiphenomenon or a proxy variable. Clearly, not every correlation

yields reliable predictions. For example, considering random sequences of finite length of a large number of binary variables, there will be arbitrarily good correlations between some of those variables, as the number of variables increases. However, these correlations are by definition completely spurious and entirely useless for prediction. For example, one of those sequences may replicate the binary representation of the number pi up to arbitrary accuracy. However, this does not guarantee predictive success as the next number in the sequence will again be completely random and in general will not coincide with the next digit of pi.

Thus, in order to guarantee even some amount of predictive success beyond mere chance, correlations are not enough because they might be spurious in the above sense. Therefore, some additional justification is required to use observed correlations for prediction. For example, the correlation could result from a direct causal relationship, either from cause to effect or from effect to cause, or from an underlying common cause structure. Scientific models of a phenomenon may also serve as additional justification, but such scientific models typically describe causal relationships, in particular if causation is understood broadly in the sense of difference making. In fact, I cannot imagine any justification for reliable prediction based on correlations that does not in some way invoke causal models—but I am happy to be proven wrong.

Based on the defence of variational induction as sketched above, in my view, the conclusion remains valid that causal models can be derived from observation statements alone and do not require experimentation. As I have argued in the past, though, experimentation has a considerable pragmatic advantage over observation in that it typically requires a controlled environment thereby largely ensuring homogeneity as a crucial requirement of variational induction. By contrast, if causal inferences are based on observation statements alone, homogeneity is in general much more challenging to establish.

References

Bacon, Francis (1620/1994). *Novum Organum*. Chicago, Il: Open Court.

Baumgartner, Michael, and Gerd Graßhoff. 2004. *Kausalität und kausales Schließen*. Norderstedt: Books on Demand.

Cohen, L. Jonathan (1989). *An Introduction to the Philosophy of Induction and Probability*. Oxford: Oxford University Press.

Galli, Serena (2023). Causal Inference from Big Data? A Reply to Pietsch (2021). *Dialectica* 77(3): 319–333.

Herschel, John F. W. (1830). *Preliminary Discourse on the Study of Natural Philosophy*. London: Longman, Rees, Orme, Brown, & Green.

Mackie, John L. (1980). *The Cement of the Universe. A Study of Causation.* Oxford: Clarendon Press.

Mill, John S. (1843). *System of Logic*. London: John W. Parker.

Pietsch, Wolfgang (2016). *A difference-making account of causation*. philsci-archive.pitt.edu/11913/.

Pietsch, Wolfgang (2021). *Big Data*. Cambridge University Press.

Pietsch, Wolfgang (2022). *On the Epistemology of Data Science. Conceptual Tools for a New Inductivism*. Springer.

Pietsch, Wolfgang (2026a). A Formal Summary of Variational Induction. www.wolfgangpietsch.de/pietsch_formal-framework.pdf

Pietsch, Wolfgang (2026b). Nothing New Under the Sun – Large Language Models and Scientific Method. philsci-archive.pitt.edu/28259/

Russo, Federica (2007). The Rationale of Variation in Methodological and Evidential Pluralism. *Philosophica* 77:97-124.

Russo, Federica (2009). *Causality and Causal Modelling in the Social Sciences. Measuring Variations.* Springer.