**Online first in: The British Journal for the Philosophy of Science**

https://www.journals.uchicago.edu/doi/10.1086/716769

# The Counterpossibles of Science Versus the Science of Counterpossibles

**Daniel Dohrn**

**Abstract**

Orthodoxy has it that all counterpossibles are vacuously true. Yet there are strong arguments both for and against the use of non-vacuous counterpossibles in metaphysics. Even more compelling evidence may be expected from science. Arguably philosophy should defer to best scientific practice. If scientific practice comes with a commitment to non-vacuous counterpossibles, this may be the decisive reason to reject semantic orthodoxy and accept non-vacuity. I critically examine various examples of the purported scientific use of non-vacuous counterpossibles and argue that they are not convincing. They neither establish that scientific practice comes with a commitment to the non-vacuity of counterpossibles, nor that incurring such a commitment would be useful in scientific practice. I illustrate a variety of counterstrategies on behalf of orthodoxy.

*1 Computability Theory – Prima Facie Reasoning From The Antecedent To The Consequent*

*2 Scientific Explanation – Interventions Made Possible*

*2.1 Difficulties with explanatory counterpossibles*

*2.2 Doing with counterfactuals*

*3 Modeling: Reasoning By Analogies Replaces Counterpossibles*

*4 Superseded Theories*

*4.1 Doing with vacuous truth*

*4.2 Counterpossibles read epistemically*

*5 Conclusion*

Philosophers have claimed that counterfactual conditionals ('counterfactuals'), usually expressed as if A were/had been the case, C would/would have been the case (formalized as A□→C) play an important role in science, for instance in formulating laws and explanations (Goodman [1947], Woodward [2003]). Yet their meaning is contentious. The Lewis–Stalnaker semantics represents orthodoxy (Stalnaker [1968], [1981], Lewis [1973]). Slightly simplifying, a counterfactual is true precisely if the consequent C is true in all closest possible worlds in which the antecedent A is true, that is in those antecedent worlds that minimally differ from the actual world. Moreover, a counterfactual utterance usually comes with some kind of presupposition that A could have been true; it is true in some possible world (Lewis [1973], p. 3, Khoo [2015], p. 21). If the presupposition fails because A could not have been true, the counterfactual ('counterpossible') is vacuously true whatever the consequent C is, comparable to a vacuous universal quantifier (all unicorns…). I shall call this position 'vacuism' and its denial 'non-vacuism' (following Berto *et al.* [2018]).[1]

The assumption that all counterpossibles are vacuously true has been questioned. Many of the dissenters are working in the field of metaphysics. However, some of the examples of non-vacuous counterpossibles presented are from science, mathematics, and logics. To give an example, the following seems false (Nolan [1997]):

[1] Possibility is determined by a contextually determined accessibility relation, yet we can disregard that refinement for my purposes and simply assume that the vacuity thesis concerns metaphysically impossible antecedents. It is notoriously difficult to tell what metaphysical impossibility is, but the examples of counterpossibles to come are uncontested instances of it.

If intuitionistic logic were correct, then the law of the excluded middle would hold.

Other examples come from science. There is a debate on the modal implications of the Lotka–Volterra model in population biology (Williamson [2017], pp. 473, 482). The model uses metaphysically impossible idealizations. One such idealization is that populations of discrete individuals (for instance wolves and rabbits) can be represented as continuous quantities. It has been suggested that these idealizations are best accounted for in terms of impossible dispositions, that is dispositions for a population to behave under impossible circumstances in which there is a continuous quantity of rabbits (Jenkins and Nolan [2012], p. 746). There is a close connection to counterpossibles about such circumstances. Encouraged by such pioneering work, several authors have more recently taken a closer look at a range of examples from science, mathematics, and logics. Their result is that the best practice in these disciplines supports the non-vacuous truth and falsity of counterpossibles. I shall critically discuss three recent positions along these lines.

In particular, Matthias Jenny [2018] argues that the best practice in computability theory as a subfield of mathematics and logics comes with a commitment to the non-vacuous truth and falsity of counterpossibles, respectively. Peter Tan [2019] makes the same point for explanation, modeling, and reasoning from superseded theories in the natural sciences. The argument for modeling is further elaborated by McLoone [forthcoming].

Both Jenny and Tan opt for philosophical humility: philosophical and linguistic analysis must respect good scientific practice (Jenny [2018], section III, Tan [2019], p. 58). Yet only actual scientific practice can claim this prerogative. It should therefore make a difference whether the reasoning (of scientists) is clearly counterfactual and the antecedent is indeed impossible, for instance because scientists themselves assert and deny counterfactuals with impossible antecedents or something akin to them, or whether a use of counterpossibles is just attributed or suggested to them by philosophers. The reconstruction of scientific practice by philosophers is just another piece of philosophical theory and should be treated accordingly.

I am undecided about the semantic issue of vacuity. My aim is only to critically weigh certain arguments against vacuity. Without aspiring to finally decide the issue of whether science comes with a commitment to the non-vacuity of counterpossibles, or at least whether such a commitment would be useful, I shall argue that the selected arguments from philosophy of science are not compelling. I shall also use the positions discussed to outline more general strategies which may be used to reap the benefit of using counterfactuals and counterpossibles in science without incurring a commitment to the non-vacuity of counterpossibles.

I add a general note of caution: some vacuists deny that counterpossibles can have an informative role (quotes in Tan [2019], p. 35). This claim should be resisted. Irrespectively of whether they are vacuous or not, counterpossibles may be used to say something informative about the world. As we shall see, their information value does not always depend on their non-vacuity, but on embedding them into the right argumentative context (Williamson [2007], p. 173, [2018], section 2, Jenkins [2010], p. 258). Thus, showing that scientists do or could make good use of counterpossibles does not yet establish that they are committed to their non-vacuity.

I give an overview of the arguments to come: (1.) I start with Jenny's argument that we need non-vacuous counterpossibles to understand reducibility in computability theory. I argue that this understanding can be better explained by a heuristic process of reasoning from the antecedent to the consequent, disregarding the impossibility of the antecedent. For Tan's example of (2.) a counterfactual account of scientific explanation, I argue (2.1.) that Tan's candidate for an explanatory counterpossible does not fit into standard counterfactual theories of explanation, even if the latter are extended such as to become hyperintensional. (2.2.) An alternative counterfactual which does with a metaphysically possible intervention is a better candidate for playing the intended explanatory role. (3.) For the case of scientific modeling, I argue against Tan and McLoone that their counterpossibles can be replaced without loss by reasoning in terms of analogies and idealizations. (4.) For the case of reasoning from superseded theories, I argue that (4.1.) such reasoning can be done by use of vacuous counterpossibles, a heuristic process as in (1.), or in terms of (4.2.) an epistemic interpretation of counterfactuals with impossible antecedents.

## 1 Computability Theory – Prima Facie Reasoning From The Antecedent To The Consequent

Jenny presents a striking example of the actual use of counterpossibles in scientific publications. He contends that computability theory is committed to reading these counterpossibles non-vacuously. Yet as I shall argue, there is a better way of understanding them. This understanding perfectly underpins their role as described by Jenny without incurring a commitment to their non-vacuous truth and falsity. It merely draws on a heuristic reasoning process that guides experts' intuitions concerning their truth and falsity.

Jenny considers the following counterpossibles:

'[Halting] If the validity problem were algorithmically decidable, then the halting problem would also be algorithmically decidable,

which is true, and

[Arithmetic] If the validity problem were algorithmically decidable, then arithmetical truth would also be algorithmically decidable,

which is false.' (Jenny [2018], p. 530])

These are counterpossibles; it is a mathematical/logical truth that the validity problem is not algorithmically decidable, and such truths are metaphysically necessary. Jenny argues that reading these counterpossibles non-vacuously is indispensable to understanding the notion of reducibility. The understanding is given by a necessary and sufficient condition. Problem A is reducible to problem B precisely if:

(Scheme) If problem B were algorithmically decidable, problem A would be.

Applying this condition, vacuity would give us the wrong result in the case of arithmetical truth: assuming vacuous truth, (Arithmetic) is true. As a consequence, the problem of arithmetical truth would reduce to the validity problem, but this is false. Only a non-vacuous reading gives us the right results. Since our understanding of reducibility depends on using (Scheme), it comes with a commitment to the non-vacuous truth and falsity of counterpossibles.

Now even if counterpossibles are indispensable in guiding our understanding of reducibility, we do not need a commitment to their non-vacuous truth or falsity. We only need an intuitive grasp of them together with a heuristic: consider whether the scientific debate at stake provides an argument that leads from the antecedent A to the consequent C without drawing on information whether the antecedent is metaphysically possible. If there is such an argument, the pertinent counterpossible A□→C seems prima facie true and (simplifing a bit) its denial A□→not-C false. These intuitions are compatible with all counterpossibles being vacuously true as the argument considered may be overridden by independent semantic reasons supporting vacuous truth.

I shall argue that a reading in terms of heuristic reasoning from the antecedent to the consequent even better fits the role of counterpossibles as described by Jenny compared to one that comes with a commitment to non-vacuity. It is not only commendable as it avoids unnecessarily strong commitments. It also better conforms to the precise role of counterpossibles as described by Jenny.

The strategy of explaining intuitions supporting the truth and falsity of certain counterpossibles by heuristic reasoning is not new. The most prominent heuristic on offer is Williamson's ([2017], [2018], [2020]). According to Williamson [2020], in assessing a counterfactual A□→C, we run an imaginative simulation. We imagine a non-actual scenario in which the antecedent A holds together with a contextually determined set of background truths and consider what attitude to take towards whether the consequent C holds in the scenario. We take the same attitude towards A□→C. If the attitude towards C is acceptance, we accept A□→C. If we accept C as true in the scenario, we reject not-C as false in the scenario. As a consequence, whenever our acceptance of C makes us accept A□→C, it makes us reject A□→not-C as false. This mostly leads to correct results, but in the case of counterpossibles, these results conflict with vacuism. According to the latter, both A□→C and A□→not-C are true when A is impossible. Williamson ([2017], [2018]) uses versions of the heuristic to debunk anti-vacuity intuitions. These intuitions can be explained by heuristic reasoning, but they are incorrect.

Williamson's use of the heuristic has been criticized (Berto *et al.* [2018], 4.2.). I want to remain neutral with regard to the general question of vacuism. Therefore my aim is not to generally debunk non-vacuity intuitions. Rather I want to show that, regardless of whether they are ultimately correct or spurious, such intuitions explain the usefulness of counterpossibles in cases of scientific reasoning like Jenny's. Williamson grants that his heuristic is generally useful. I shift the emphasis in claiming that his even holds for non-vacuity intuitions. Yet in order for intuitions to be useful in a scientific debate, the reasoning that guides them should be subject to certain qualitative standards. It should conform to the methods and standards that are prevalent in the debate. Though I shall not formulate the reasoning in terms of Williamson's imaginative process, I see no obstacle for doing so.

To support my claim, I build on a more positive view of the role of prima facie reasoning in terms of counterpossibles as sketched by Carrie Jenkins:

‘Even a proponent of the traditional semantics can appreciate that some counterpossible conditionals are non-trivially interesting as well as trivially true. For instance, certain counterpossible conditionals can appear at the beginning of persuasive reductio proofs (e.g. “[Number] If √2 were rational it could be written as n/m with n, m integers”), because there are non-trivial reasons for taking these counterpossibles to be true — reasons which can be accepted by someone who does not already accept that their antecedents are impossible.’ (Jenkins [2010], p. 258)

Jenkins uses a mathematical example reminiscient of Jenny’s. Certain counterpossibles may be interesting, play a role in inquiry, even if they are vacuously true. There are reasons for (or against) their truth which are independent of the metaphysical impossibility of the antecedent. The working criterion of their independence is that they are acceptable to someone who does not yet accept the impossibility of the antecedent or simply disregards the question of whether the antecedent is possible in the argument. For instance, one may argue for (Number) as follows:

(Number1) Antecedent: √2 is a rational number.

(Number2) Independently of whether the antecedent is true or false, (Number3) holds:

(Number3) Any rational number can be written as n/m with n, m integers.

(Number4) Consequent: √2 can be written as n/m with n, m integers.

I add (Number2) as a test for the contenability of the other premisses with the antecedent.

In order to explain intuitions regarding the truth of a counterpossible, I suggest the following: we prima facie accept a counterfactual A□→C if there is an argument from A to C, disregarding whether A is possible. We prima facie reject A□→C if (i) there is such an argument from A to not-C, and (ii) there is no such argument from A to C.[2] The arguments should live up to the standards of the scientific debate at stake.

In this vein, I propose the following explanation of the intuitions reported about (Halting) and (Arithmetic): computability theorists prima facie accept (Halting) because there is an argument conforming to the standards of computability science that leads from the antecedent to the consequent without drawing on information whether the antecedent is possible. The argument is the following:

(Validity) Antecedent: The validity problem is algorithmically decidable.

(Halting1) Independently of whether the antecedent is true or not, (Halting2) holds:

(Halting2) There is a set of algorithmic transformations such that, if solutions to the validity problem are given, they are turned into solutions to the halting problem.

(Halting3) Consequent: The halting problem is algorithmically decidable.

In a similar vein, theorists tend to reject (Arithmetic) because an argument conforming to the standards of computability science leads them from the antecedent to the denial of the consequent without drawing on information whether the antecedent is possible, and there is no parallel argument from the antecedent to the consequent:

---

[2] The condition is sufficient but not necessary. I need clause (ii) to evade a counterexample to Williamson's heuristic (Berto *et al.* [2018], p. 707). Both of the following seem true:

If it were raining and not raining, it would be raining.

If it were raining and not raining, it would not be raining.

(Validity) Antecedent: The validity problem is algorithmically decidable.

(Arithmetic1) Independently of whether the antecedent is true or not, (Arithmetic2) holds:

(Arithmetic2) There is no set of algorithmic transformations such that, if solutions to some other problem are given, they are turned into solutions to the problem of arithmetic truth.[3]

(Arithmetic3) Denial of the consequent: It is not the case that the problem of arithmetic truth is algorithmically decidable (see Jenny [2018], p. 534).

To show that we can do with intuitions based on such a heuristic reasoning, I summarize the role of (Halting) and (Arithmetic) in relative computability theory as described by Jenny: in a first step, it is independently established that the validity problem, the halting problem, and arithmetical truth are not algorithmically decidable. To show this, abstract machines called Turing machines are introduced. For each problem, it is shown that Turing machines cannot solve it. Assuming the Church–Turing thesis that Turing machines are an adequate model for algorithmic decidability, it follows that the problems are not algorithmically decidable (Jenny [2018], p. 533).

The reducibilty of the halting to the validity problem and the irreducibility of the problem of arithmetical truth to the validity problem are established in a parallel manner. In a first step, 'oracle Turing machines' are introduced. These are Turing machines enhanced by an additional storage device (the oracle), which provides the answers to a problem B. A problem A is Turing–reducible to B precisely if the answers to B can be transformed into answers to A.

[3] This assumption is formulated such as to rule out the alternative that arithmetic truth is decidable independently of whether the validity problem is, which would threaten any use of (Arithmetic) as a criterion of reducibility.

In a second step, the post–Turing thesis or relativized Turing thesis is introduced. The thesis says that A is reducible to B just in case it is Turing–reducible (Jenny [2018], p. 534).

Counterpossibles enter the picture at the second step: in order to assess the post–Turing thesis, we need an independent understanding of reducibility. The role of this understanding is described by Jenny as follows, deferring to the standard textbook by Rogers [1967]:

‘[Rogers] produces two sets, the first of which he argues is reducible to the second. He then shows that the first set isn’t truth-table reducible to the second, but that it is Turing reducible to it. Rogers concludes that using truth-table reducibility to analyse what he explicitly calls the intuitive notion of reducibility would be inadequate, for this would leave out certain sets, and that an analysis in terms of Turing reducibility fares better. To arrive at this verdict, Rogers clearly assumes that he and his readers have an understanding of the notion of reducibility that’s independent of talk about oracle Turing machines. And the understanding of reducibility that Rogers provides is in terms of counterfactuals.’ (Jenny [2018], p. 544)

Jenny refers to the following passage from Rogers:

‘Intuitively, A is reducible to B if, given any method for calculating [the characteristic function of B], we could then obtain a method for calculating [the characteristic function of A.]’ (Rogers [1967], p. 127)

The interpretation of this key quote is decisive. There are two uncertainties. First, Rogers does not explicitly deny counterpossibles like (Arithmetic).[4] Second, Roger’s use of ‘could’ reminds

[4] ‘[…] none of the examples of negated mathematical counterfactuals in Jenny (2018) are drawn from the writing of actual mathematicians.’ (Yli-Vakkuri and Hawthorne [2020], p. 575). Yli-

us of a ‘might’- instead of a ‘would’-counterfactual. However, in Lewis’s analysis, all ‘might’-counterfactuals with impossible antecedents are false ([1973], p. 21). One may resort to an epistemic interpretation as ‘perhaps we would obtain a method for calculating’ (Stalnaker [1981]), but this seems too weak to support Rogers’s argument. He should rather be read as saying ‘would then be able to obtain.’[5]

Notwithstanding these doubts, I follow Jenny in assuming that the intuitions invoked by Rogers concern the truth and falsity of counterpossibles instantiating (Scheme). However, I suggest a cautious reading of ‘intuitively’ as ‘there is an intuition that…’, or ‘it seems that…’, which does not require the intuitions to be correct. The cautious reading may be put either as an interpretation of or as an amendment to Rogers’s position.

In Rogers’s argument as reconstructed by Jenny, different alternatives for analysing the notion of reducibility are tested against intuitive cases of reducibility. The cases are described by applying (Scheme): if problem B were decidable, A would be. Counterpossibles like (Halting) and (Arithmetic) provide an intuitive classification of cases of reducibility and irreducibility. Competing analyses of the notion of reducibility are tested as to whether they support this classification. The alternative that supports the classification is proposed as an analysis.

---

Vakkuri and Hawthorne also criticize that counterpossibles do not play a role in a formal proof but that their role is confined to providing a first intuitive criterion of reducibility. My argument targets this limited role of counterpossibles.

[5] Thanks to an anonymous reviewer for this suggestion, which is also supported by other formulations of the same idea like the following: ‘*The solutions to which problems would also furnish solutions to P*?’ (Davis [1958], p. 179, emphasis in the original), as quoted in Jenny ([2018], p. 531).

I note two remarkable features of this procedure: first, we do not need the actual truth and falsity of counterpossibles like (Halting) and (Arithmetic) to arrive at the intuitive classification. We only need relevant intuitions about their truth and falsity, respectively. Second, Jenny's talk of an analysis is evidence that Turing–reducibility then replaces our intuitions by a more precise understanding of reducibility.

For our intuitions to be relevant in computability theory, they should be backed by an argument that leads from the antecedent to accepting or denying the consequent. In the case of (Halting) and (Arithmetic), the argument is that there is an algorithmic procedure that leads from given solutions to the validity problem to solutions to the halting problem, but there is no comparable procedure for arithmetical truth. I have encountered the objection that our intuitions cannot provide a success criterion for a definition of reducibility unless they are correct. Yet they do not have to be correct as far as they lead to assessing certain counterpossibles as non-vacuously true and false. It is only required that they support a correct classification of cases.[6]

To arrive at a correct classification, we do not need to address whether the antecedent of the pertinent counterfactuals is possible. On the contrary, there are reasons to assume that our considerations circumvent the issue of whether the antecedent is possible. Consider how the notion of Turing–reducibility is attained. The very rationale of introducing oracle Turing machines is to get around the issue of algorithmic decidability. Regardless of whether a problem B is algorithmically decidable, the oracle is supposed to give us the results which otherwise

[6] One may compare the role of our intuitions about counterpossibles to that of folk physics. Though being false, it forms part of a reliable mechanism that supports many correct predictions in everyday matters (Williamson [2007], p. 145). I also note that, instead of a wrong intuition about (Arithmetic), we may do with a correct intuition that, if the validity problem were algorithmically decidable, then arithmetical truth would still not be algorithmically decidable (see the discussion on modeling water in section 3).

would have to be attained by algorithmically deciding the problem. We do not have to bother any longer about whether problem B is decidable, that is whether the antecedent of (Scheme) is possible or not.

It seems plausible that the implicit considerations that guide our intuitions already anticipate the introduction of oracle Turing machines. The guiding idea is the following: assuming we somehow get the answers to problem B, is there a way of transforming them into answers to problem A? Oracle Turing machines make this idea precise.[7]

The reasons which guide our intuitions with regard to counterpossibles like (Halting) and (Arithmetic) can be put as follows: we consider whether a solution to problem B can be transformed into a solution to problem A. We conspicuously disregard whether there could be an algorithmic procedure of solving problem B. Since our reasons do not address –but are selected such as to avoid addressing– the possibility or impossibility of the antecedent, it seems more appropriate to interpret our intuitive understanding of reducibility as not involving a commitment to non-vacuous truth and falsity of counterpossibles.

In sum, often we may make do with intuitions on counterpossibles and their prima facie reasons instead of a commitment to non-vacuous truth or falsity.

## 2 Scientific Explanation – Interventions Made Possible

Coming to Tan's [2019] arguments, I shall use them to illustrate several strategies for avoiding a commitment to non-vacuity. The first is to show that the purported role of a counterpossible

[7] One may suspect that the argument already presupposes an understanding of reducibility by something akin to the post-Turing thesis. Yet the decisive point is that we do not yet make use of the notion of an oracle Turing machine (Jenny [2018], p. 544). The explicit post–Turing thesis may be interpreted as explicating our implicit understanding.

is more convincingly played by a counterfactual with a metaphysically possible antecedent. In the case at hand, this role is scientific explanation.

Here is Tan's main example:

'(D) If diamond had not been covalently bonded, then it would have been a better electrical conductor.

The macroscopic property of diamond's poor electrical conductivity is brought about by, and hence depends on its microstructural bonding properties. This counterfactual conditional, (D), is indispensable to providing a scientific explanation of why diamond has that macroscopic property, since it identifies and describes that dependence relation.' (Tan [2019], p. 40–1)

Tan argues that (D) is a genuine counterpossible. For diamond not to be covalently bonded, it would have to lack the very atomic structure that makes it diamond. Moreover, at first glance, (D) seems true and its denial (would not have been a better conductor) false. However, such intuitions on (D) in themselves do not add much to the many other non-vacuist intuitions proferred by philosophers. The special status of intuitions on (D) as evidence for non-vacuism arises from their purported indispensability for scientific explanation.

As noted initially, we are faced with a philosophical reconstruction rather than first-hand scientific practice. (D) is not directly taken from actual scientific discourse. Yet Tan's claim that some particular counterfactual plays the core role in explaining the poor conductivity of diamond is supported by an influential counterfactual view of explanation. According to such a view, any explanation depends on specifying some perfectly corresponding counterfactuals.

I shall proceed in two steps. In (2.1.), I shall point out difficulties for the counterpossible (D) to fit the counterfactual theory of explanation as invoked by Tan. In (2.2.), I shall present a counterfactual conditional with a bona fide possible antecedent that better fits the explanatory role according to such a theory.

### 2.1 Difficulties with explanatory counterpossibles

Tan's claim depends on the role of (D) in a convincing overall account of explanation. Tan uses Woodward's [2003] theory. It therefore seems legitimate to judge Tan's proposal by this theory. Woodward defines causality by possible interventions. The challenge posed by Woodward's account for Tan becomes the following: can the admittedly metaphysically impossible antecedent of (D) still be the subject of a possible intervention in any useful sense?

In section (3.5.) of Making Things Happen, Woodward discusses the notion of possibility at stake, asking: 'Must interventions be physically (i.e. nomologically) possible for X to cause Y?' (Woodward [2003], p. 128) Woodward answers this question in the negative.[8] Yet in what sense do interventions have to be possible? Woodward gives a characterization of the possibility at stake:

'[…] the reference to "possible" interventions […] does not mean "physically possible"; instead, an intervention on X with respect to Y will be "possible" as long as it is logically or conceptually possible for an intervention on X with respect to Y to occur […] The sorts of counterfactuals that cannot be legitimately used to elucidate the meaning of causal claims will

[8] An anonymous reviewer has suggested that for Woodward metaphysical possibility patterns with nomological possibility. One might therefore read him as claiming that interventions do not have to be metaphysically possible. Yet as we will see, it is still unlikely that Woodward accepts interventions that would undo covalent binding while preserving diamond.

be those for which we cannot coherently describe what it would be like for the relevant intervention to occur at all […]' (Woodward [2003], p. 132)

A direct intervention on the cause C, that is a process directly changing C without thereby directly changing anything that is not causally downstream from C, must be possible in a sense that Woodward characterizes by 'conceptually', 'logically', 'analytically'. Moreover, we must be able to describe the intervention coherently.

Woodward's condition that an intervention must lend itself to being coherently described does not square with Tan's explicit interpretation of the antecedent of (D). According to Tan, it is 'straightforwardly contradictory' to assume anything to be diamond without being covalently bonded (Tan [2019], p. 41). Tan does not tell how the contradiction arises. One salient interpretation treats (D) akin to statements like 'water is not $H_2O$', following textbook Kripkeanism.

The contradiction is not a priori, and it is not explicit, but it may be made explicit as follows: we have to assume something akin to a property identity between the property of being diamond and the property of being X and covalently bonded, where 'X' stands for the rest of diamond's essential features. The intervention Tan has in mind is one that makes an item that is and remains diamond not covalently bonded. Assuming referential transparency, we can replace the property of being diamond by the property of being X and covalently bonded.[9] Thus, an intervention that results in diamond without covalent bonding has to be described as resulting in the same item being covalently bonded and not covalently bonded. The intervention cannot

[9] Non-vacuists will balk at the assumption of referential transparency. Yet at this point I am trying to make sense of Tan's claim that it is contradictory to suppose anything to be diamond without being covalently bonded. If the best way of doing so conflicts with non-vacuism, the worse for Tan.

be coherently described.[10] There is no way for an intervention to make the same item at the same time both covalently bonded and not covalently bonded. One may try to conceive of a situation in which the same item is both covalently bonded and not covalently bonded in a discussion of dialetheist logic, but such considerations have no place in the science of conductivity.

This diagnosis is confirmed by Woodward's own discussion of surgical interventions in the case of relationships of definition and supervenience. According to Woodward [2015], an intervention on variables that are related by definition and supervenience has to leave these relations intact. For instance, there can be no intervention on a mental property which is not at the same time an intervention on its supervenience base and vice versa.[11] It seems highly plausible that Woodward would claim the same for the relationship between being diamond and the underlying property of being covalently bonded. Put in terms of causal modeling, there is no intervention that sets the variable being covalently bonded to 0 and keeps the variable being diamond at 1.

As it stands, Tan's argument that scientific explanation commits us to the non-vacuity of (D) is unsatisfying. One may consider the option of modifying interventionism as traditionally conceived by Woodward and others such as to attain a hyperintensional approach to explanation. One exemplary approach along these lines has been presented by Wilson [2018] for grounding relationships. Take the standard example of Socrates and Singleton Socrates, the set {Socrates} with Socrates as its only member (Wilson [2018], p. 721–2). Necessarily, Socrates exists if and only if {Socrates} exists. However, the grounding relationship is

[10] It may be coherently described at a superficial level, but it surely should be coherent 'all the way down', giving more informative structural descriptions.

[11] Of course, there may be interventions on the supervenience base that do not touch its subvening role.

purported to go one way only. The existence of Socrates grounds the existence of {Socrates}, but not vice versa. According to Wilson, this grounding relationship can be described by a model that is analogous to a causal model, specifying variables (C: whether Socrates exists; E: whether {Socrates} exists), Structural equations (C=E), assignments (C=1; E=1) and graphs that do not represent causal but grounding relationships (C=1 → E=1). One may then conceive of an intervention that surgically undoes whether Socrates exists (C), and an intervention that surgically undoes whether {Socrates} exists (E). The results of the interventions can be represented by non-vacuous counterpossibles:

(Socrates1) If an intervention had prevented Socrates from existing, then {Socrates} would not have existed. — True

(Socrates2) If an intervention had prevented Socrates from existing, then {Socrates} would have existed. — False

(Socrates3) If an intervention had prevented {Socrates} from existing, then Socrates would not have existed. — False

(Socrates4) If an intervention had prevented {Socrates} from existing, then Socrates would have existed. — True

I don't take stance on the plausibility of such an extension of the interventionist account to grounding relationships and the underlying metaphysical assumptions, but I note that it still does not fit Tan's needs. Tan would need to conceive of an intervention that undoes covalent bonding while preserving diamond. If a relationship of metaphysical priority obtains between being diamond and covalent bonding, it goes from being covalently bonded to being diamond. One may say that being covalently bonded partially grounds or constitutes being diamond. Yet as a consequence, any metaphysical intervention as envisaged by Wilson that would undo covalent bonding would also undo being diamond, just as an intervention that would have prevented Socrates from existing would have prevented {Socrates} from existing.

Summarizing, even a daring metaphysical extension of a counterfactual account of explanation does not support Tan's use of (D) in explaining the poor conductivity of diamond.

### 2.2 Doing with counterfactuals

In the last sections, I have raised doubts that (D) as read by Tan can figure in an explanatory relationship. I shall now propose to replace (D) by a different counterfactual that arguably is not a counterpossible. I claim that this counterfactual identifies the explanatory relationship as effectively and better fits into a counterfactual account of explanation as developed by Woodward.

We have to explain the weak conductivity of diamond by covalent bonding. The relevant intervention should satisfy several requirements. Since it is a property of diamond that is to be explained, namely that diamond is a poor conductor, diamond must figure in the explanation. It must also figure in the counterfactual that is used in the explanation. One may argue that diamond must be preserved throughout the intervention to ensure that it is the property of diamond and nothing else that is explained. This requirement is overdemanding. To be sure, we need diamond as the item whose properties are to be explained. Yet it suffices that the intervention operates on diamond. It is not required that it preserves diamond.

Before the intervention, diamond has the property of being a poor conductor, and that property gets lost when the intervention surgically undoes covalent bonding. Since the intervention is surgical, it leaves other features of diamond untouched as far as possible, for instance the composing structure of carbon atoms. We may think of the intervention as operating only on parts of the electron shell. The result is that the conductivity of diamond disappears while other features that partially constitute diamond remain. This result allows us to pinpoint the explanans of the original conductivity property to the property removed by the intervention. One may doubt that even such an intervention is metaphysically possible, but here the burden of proof clearly would be on the side of the opponent. The intervention seems bona fide possible.

How do we ensure that the intervention is one on diamond? In the original (D), diamond is addressed generically. We may preserve this generic way of speaking of diamond, but we need to avoid the contradiction that arises when we characterize the result of the intervention as diamond. One way of avoiding it is to treat being diamond as a property of something that does not necessarily have this property. I suggest that we talk of the matter that is actually diamond. 'Matter' is a standard term in physics. Matter can instantiate different properties like being diamond, being iron, and so on. Matter can become diamond, and it can cease to be diamond. Thus, it can act as a persistent referent that survives the intervention while diamond does not.

(D*) For any matter m that is diamond, if there had been an intervention undoing the covalent binding of m, m would have been a better conductor.

One may demur that (D*) does not figure in explaining the poor conductivity of diamond rather than the matter that is diamond. Yet our talk of diamond being a poor conductor is vague. Nothing seems to be lost by rephrasing it in the way proposed. One may even consider (D*) as one legitimate way of making the everyday language (D) precise in an explanatory context.

In sum, even if we subscribe to a counterfactual-based account of explanation, there are relevant alternatives to incurring a commitment to the non-vacuity of counterpossibles. The strategy explored in this section was to circumvent the impossible intervention as described in (D) by a possible intervention as described in the counterfactual (D*). It is not guaranteed that this can always be done, but in the case of (D), it seems perfectly feasible.[12] As long as these alternatives are not exhausted, it has not been established that scientific explanation comes with a commitment to non-vacuous counterpossibles.

## 3 Modeling: Reasoning By Analogies Replaces Counterpossibles

I shall come to another argument of Tan's, which I shall use to illustrate a different debunking strategy. Often a counterpossible can be replaced by spelling out relevant (dis)analogies between reality and an antecedent scenario.

Scientists use models for different purposes, for instance to represent overcomplicated situations. How exactly to account for this practice is contentious. Tan claims that the practice of modelling is entangled with a commitment to non-vacuous counterpossibles. He quotes many philosophers accounting for models in terms of fictions, counterparts, and counterfactuals, yet none of them addresses counterpossibles and the issue of vacuity.[13] Tan's first example of a counterpossible in modeling is Ernest Adams on treating the planets as mass points:

'The orbits of the planets are similar to those of mass points and
[Planet] if they were mass points then their orbits would be such-and-such;
therefore their orbits are similar to such-and-such.' (Tan [2019], p. 45, Adams [1993], p. 5)

[12] Baron *et al.* [2017] present explanations where the strategy may not work, but they do not purport to reconstruct actual scientific practice.

[13] Among the philosophers quoted as invoking counterfactuals are Bokulich [2011] and Psillos [2011].

All the examples to be considered are again from the philosophical literature, not from scientific discourse. Their special evidential value compared to other examples is purported to lie in their relevance for scientific practice. Yet as long as we do not have a concrete example of scientists asserting an explicit counterpossible (or something akin to it) in a key argument from the behaviour of models to the behaviour of target systems, we again are faced with a philosophical reconstruction, which has to be weighed against alternative reconstructions. Even if a scientist were to assert a counterpossible like Adams's, we should check whether the assertion is not simply a fancy way of gesturing at an argument that does without counterpossibles.

Given the scant direct evidence that scientists do use counterpossibles in modeling, we would expect strong arguments for why they should be reconstructed as using non-vacuous counterpossibles, or why it would be useful to use them. Tan just contends:

'There is widespread agreement that idealized models are informative when the observed behavior of their target objects closely approximates, at a counterfactual level, the idealized objects' behavior. Why can we manage to learn anything from an ideal, massless-string model of a pendulum? It is because the trajectory of an actual pendulum is quite close to what its trajectory would be like, if its string were massless.' (Tan [2019], p. 44)

Tan contends without giving further reasons that we can learn from a model because (and provided) the actual behaviour of its target objects is close to the behaviour of objects under the counterfactual supposition that the idealizing assumptions obtain. This claim in terms of counterfactual closeness states only a condition for learning from a model, but it remains silent about whether to use counterfactuals in reasoning by the model. Perhaps there are better ways for reasoning by a model that meets the closeness requirement. Moreover, nothing Tan says here on counterfactuals bears on the question of whether the use of impossible models should come with a commitment to the non-vacuity of counterpossibles.

Over and above the insufficiency of Tan's motivational considerations, there are more positive reasons for doubting them. Closeness is usually spelled out in terms of similarity. Thus, one salient alternative to using counterfactuals is to state relevant analogies and disanalogies, as it has been suggested already in the earlier literature on modeling (Hesse [1963]). The premise and the conclusion of Adam's condensed argument are already put in terms of similarity:

(Planet1) The orbits of the planets are similar to those of mass points.

(Planet) If the planets were mass points, their orbits would be such-and-such.

(Planet2) The orbits of the planets are similar to such-and-such.

In this argument, a step from the counterfactual situation to the actual situation is missing. I suggest an amendment:

(Planet) If the planets were mass points, their orbits would be such-and-such.

(Planet3) If the planets were mass points, their orbits would be like their actual orbits.

(Planet2) The orbits of the planets are similar to such-and-such.

Adam's use of a counterpossible in such an argument can be replaced without a loss of argumentative force, accuracy, or expressive power by saying something like the following:

(Planet4) Description: a set of mass points follows such-and-such orbits.

(Planet5) The orbits of the planets are as in (Planet4).

(Planet2) The orbits of the planets are similar to such-and-such.

On may insist that this argument cannot replace the use of counterfactuals as the modal status of what is described has to be clarified in order to draw a comparison. While the request seems legitimate, so far no reason has been given that non-vacuous counterpossibles provide a good answer to it, or that science has to come with a commitment on how to answer it. As long as such reasons are not forthcoming, one may simply leave the modal status of the model open. Grasping the condition for a sufficiently comprehensive description as in (Planet4) to be true suffices for drawing certain comparisons.

The advantage of the statement in terms of similarity is that it can be made ever more precise by specifying relevant commonalities and differences between the model and the target system. The model posits point masses, whereas the target system has bulky planets, but the law of gravity is the same, and so on. As long as we are not told what is to be gained by hypothetically identifying planets with point masses, there is no reason to prefer a reconstruction in terms of counterpossibles to an alternative in terms of similarity and dissimilarity.

Besides the quote from Adams, Tan presents an example of his own, characteristically again a merely hypothetical case, inspired by but detached from scientific practice:

‘Suppose a group of scientists constructs a pair of mathematical models of some water, M1 and M2, that each represent it as a continuous, ideal fluid, but differ with regard to the viscosity they ascribe to it. These scientists are, let us suppose, interested in how waves behave when they crash up against a certain kind of rough surface. In testing these models, what they will do is closely observe how water actually behaves when its waves crash against such a surface, and check the models’ predictions against those observations. Imagine that M1 is found to closely approximate the observed behavior, while M2’s predictions are a bit farther off. In rejecting M2 in favor of (tentatively) accepting M1, the scientists will thus take

[WaterM2] "If water were a continuous, incompressible medium, then it would behave as M2 predicts"

to be false, while simultaneously taking

[WaterM1] "If water were a continuous, incompressible medium, then it would behave as M1 predicts"

to be true.' (Tan [2019], p. 47)

Tan contrasts the true counterpossible (WaterM1) and the false (WaterM2) in order to distinguish correct and incorrect predictions from competing models.

Coming to my discussion, a reviewer has suggested that, instead of the falsity of (WaterM2), we could do with the truth of

(WaterNM2) If water were a continuous, incompressible medium, then it would not behave as M2 predicts.

To decide the issue, we should consider whether such a replacement would dispense with non-vacuity. After all, vacuism is also committed to the truth of

(WaterM2) If water were a continuous, incompressible medium, then it would behave as M2 predicts.

(WaterNM1) If water were a continuous, incompressible medium, then it would not behave as M1 predicts.

To use vacuously true counterpossibles, we would therefore need a criterion to decide between those that may and those that may not be used in arguing from the model. Again one may choose the strategy of distinguishing between interesting and uninteresting counterpossibles outlined in section (1.).[14] We may use a counterpossible provided there is an argument that leads from the antecedent to the consequent without drawing on information as to whether the antecedent is possible. The argument should conform to the standards of the scientific debate at stake.

(Water1) Antecedent: Water is a continuous, incompressible medium.

(Water2) Independently of whether the antecedent is true or not, (Water3)–(Water4) hold:

(Water3) Consequent: Water behaves as predicted by M1.

(Water4) Consequent: Water does not behave as predicted by M2.

(Water2)–(Water4) are supported by empirical evidence and reasoning. The argument supports (WaterM1) and (WaterNM2) against (WaterM2) and (WaterNM1). However, the question remains what this method of selecting counterpossibles does for us. In the case of Jenny, it was used to support an intuitive classification. Yet can we make do with mere intuitions in the case of modeling? The answer depends on the overall role of models and counterpossibles in modeling. Since Tan leaves that role largely unspecified, he cannot rule out that we can do with the (vacuous) truth of (WaterNM2) instead of the falsity of (WaterM2).

This brings me to my main criticism: Tan does not give a complete story about the role of counterpossibles. Whereas (Planet) was used to reason from a model to reality, (WaterM1) and (WaterM2) seem to be used to reason from reality to a model. We are left in the dark about how to embed such limited pieces of reasoning into a general view of modeling. Since such an embedding would be crucial to appreciating the claim that modeling comes with a commitment to non-vacuity, the claim is insufficiently supported.

Even the partial argument from reality to a model is not beyond doubt. The parallel to (Planet) makes the following reconstruction salient:

---

[14] Perhaps such a strategy could also be used for Jenny’s case from section (1.).

(Water3) Water behaves as M1 predicts.

(Water5) If water were a continuous, incompressible medium, it would behave like actual water.

(WaterM1) If water were a continuous, incompressible medium, then it would behave as M1 predicts.

Analogously for the falsity of (WaterM2). Just as Adam's argument, this reasoning may be replaced by a purely similarity-based one, comparing the actual behaviour of water to that of a continuous, incompressible medium.

My doubts also apply to a more recent argument by McLoone [forthcoming]. McLoone denies that non-vacuous counterpossibles have to be used in modeling. He just claims that it is 'permissible' to do so. Such a permission may well depend on non-vacuism rather than supporting it. If we are interested in the contribution of science to the semantic issue between vacuism and non-vacuism, we have to ask what particular role non-vacuous counterpossibles can play in modeling. As long as this question is not answered, their use may be too casual for the practice of modeling qua scientific practice to support their non-vacuity. With the exception of Tan and Jenny, the literature on the use of counterfactuals in modeling as quoted by McLoone does not support non-vacuism.[15]

[15] McLoone quotes Godfrey-Smith [2020] and Williamson [2020b]. Godfrey-Smith does not incur a commitment to counterpossibles. On the contrary, he conspicuously chooses a formulation which avoids such a commitment: 'If a pair of populations had features F, then it would have/ do G (Lotka–Volterra behaviors).' (Godfrey-Smith [2020], p. 168) Since the populations are not specified, they could be continuous. McLoone does not tell us why to prefer counterpossibles. A cautious formulation à la Godfrey-Smith may also be preferable in Tan's examples. Likewise, in his own theory of models, Williamson avoids formulating models by

Drawing on the discussion in Jenkins and Nolan [2012], McLoone points out that population dynamics as in the Lotka–Volterra model of the predator-prey-relationship often uses metaphysically impossible idealizations. One such idealization is that populations of discrete individuals (for instance wolves and rabbits) can be represented as continuous quantities. In this vein, McLoone proposes that one may model population dynamics by accepting (6) and rejecting (7):[16]

‘(6) If a population of rabbits satisfied the assumptions of the logistic equation, then its size would eventually equal the carrying capacity.
(7) If a population of rabbits satisfied the assumptions of the logistic equation, then it would eventually go extinct.’

The antecedent of (6) and (7) invites us to suppose that the assumptions involved in applying the logistic equation are satisfied. An exemplary impossible assumption is that the population size is a continuous quantity. Later in McLoone’s reasoning, it transpires that the antecedent should not only include the idealizing assumptions but the concrete mathematical formulation

---

subjunctive counterpossibles, using instead indicative conditionals: ‘When we explore a model by valid deductive reasoning from the model description, we learn necessary truths of the general conditional form, “If a given case satisfies the model description, then it satisfies this other description too.”’ ([2017b], p. 162) Later he indeed considers using counterfactuals like ‘If there were any simple pendula, they would move sinusoidally’ [2020b]. Yet Williamson consistently maintains that any counterpossibles must be read vacuously ([2007], [2017], [2018], [2020]).

[16] I adopt McLoone’s numbering.

of the logistic equation. The consequent is to be derived from the antecedent by mathematical reasoning.

McLoone proposes an extended closest-worlds semantics, including both possible and impossible worlds, which is to support that (6) is true and (7) false. While the closest antecedent worlds are metaphysically impossible, an antecedent world in which the consequent of (6) is true and the consequent of (7) is false is closer than any antecedent world in which the consequent of (6) is false and the consequent of (7) is true. The reason is that the former world breaks just some actual laws of metaphysics (for instance that rabbits are discrete), but the latter additionally breaks the actual laws of mathematics.[17] Here the relevance of writing the logistic equation into the antecedent becomes crucial. The consequent is to be derived by mathematical reasoning from the antecedent.[18]

I grant that McLoone's counterpossibles are as intuitively true and false as other examples presented by non-vacuists. However, as they stand, they do not add new evidence for non-vacuism. Things would be different if their significance to science could be established. Yet McLoone does not tell what their special significance is. He does not tell what the advantage of committing oneself to (6) and the falsity of (7) is compared to the following sober statement: the consequent of (6) follows by mathematical reasoning from premise A, B, C… (those that go into the antecedent supposition), the consequent of (7) does not.

McLoone suggests that the consequent states the model's predictions. The predictions we are interested in concern the actual world. Yet a counterpossible like (6) does not directly say anything about the real world as far as the antecedent is false (and even impossible). In order to arrive at a connection to the actual world, one may again resort to analogical reasoning:

---

[17] The formulation is intended as covering both cases where the laws are violated and cases where they are not laws as distinguished by McLoone.

[18] For doubts about this claim see Godfrey-Smith ([2020], p. 171).

(6) If a population of rabbits satisfied the assumptions of the logistic equation, then its size would eventually equal the carrying capacity.

(Rabbit1) If a population of rabbits satisfied the assumptions of the logistic equation, then its dynamics would be like the actual dynamics of rabbits.

(Rabbit2) Actual rabbit populations tend towards eventually equalling the carrying capacity.

While such a reconstruction may work, there is again the relevant alternative of purely similarity-based reasoning:

(Rabbit3) Description: a population of rabbits satisfies the assumptions of the logistic equation.

(Rabbit4) By mathematical reasoning from (Rabbit3): The size of the population eventually equals the carrying capacity.

(Rabbit5) The dynamics of actual rabbit populations is as in (Rabbit3)–(Rabbit4).

(Rabbit2) Actual rabbit populations tend towards eventually equalling the carrying capacity.

Again that reasoning could replace the use of (6).

There is a further uncertainty about McLoone's counterpossibles. They differ from Tan's examples. Whereas in Tan's examples, only the idealizing assumptions are written into the antecedent and other assumptions of the models like the orbits of the point masses or M1/M2 are written into the consequent, in McLoone's case 'all of the model's assumptions' are written into the antecedent. The consequent only is thought to specify mathematical consequences of these assumptions. While there is uncertainty about how to tell apart idealizing assumptions, 'all of the model's assumptions', and consequences of a model, there seems to be a significant difference. A relevant alternative to McLoone's (6) would seem closer to Tan's examples:

(6*) If a population satisfied the idealizing assumptions A, B, C…, it would (be correctly represented by the logistic equation and thus) eventually equal the carrying capacity.

The difference between Tan's and McLoone's accounts manifests further uncertainties on the purported function of counterpossibles. One may say that Tan and McLoone put counterpossibles to different legitimate uses. Yet counterpossibles like (6) and (6*) rather seem to compete for how to use counterpossibles in modeling. Again the lack of an integrated account of the overall role of counterpossibles in formulating a model and reasoning from it raises doubts about the claim that science supports non-vacuity.

Coming to a general assessment, Tan and McLoone do not tell us what advantage a commitment to non-vacuism has in modeling compared to alternatives that avoid such a commitment. A convincing account of counterpossibles in modeling should do more than merely request that the behaviour of the target objects and the behaviour of the objects under the counterfactual supposition be close to each other. It should tell what the surplus value of using counterpossibles in modeling is.

I do not deny that counterpossibles might do something for us that cannot as well be done by reasoning in terms of relevant similarities and dissimilarities, but as long as the special role of counterpossibles has not been elucidated, and as long as there are no uncontentious examples of a use of non-vacuous counterpossibles outside of philosophical reconstructions, we have not been given good reasons for acknowledging that the scientific practice of modeling supports non-vacuism.

## 4 Superseded Theories

### 4.1 Doing with vacuous truth

The last argument of Tan's to be considered concerns reasoning from superseded theories. Often we reason by supposing theories to be true although we know them to be necessarily false. While I have questioned whether Tan's other examples are part of scientific practice at all, Tan's third example might indeed have a place in it. Still I shall illustrate two complementary counterstrategies. The first (4.1.) is to use vacuous counterpossibles for reasoning from superseded theories. This counterstrategy is highly relevant as many authors take reductio ad absurdum for instance in mathematics to depend on the vacous truth of counterpossibles (Yli-Vakkuri and Hawthorne [2020], p. 568, Williamson [2018], section 3).[19]

My first counterstrategy can be illustrated by Tan's key example, the argument:

'(B1) If Bohr's theory of the atom had been true, then an electron's angular momentum L in the ground state would have been observed at L=ħ (i.e. the Planck constant).

[19] Non-vacuists tend to disagree (Berto *et al.* [2018], 3.3). Yli-Vakkuri and Hawthorne admit that there is a tendency among mathematicians and logicians to deny certain counterpossibles. The counterpossibles they have in mind correspond to the example of reasoning from superseded theories as provided by Tan. Yli-Vakkuri and Hawthorne suggest that the tendency might be explained by the use of heuristics of the sort developed by Williamson. In my discussion of Williamson's heuristic in section (1.), I have already pointed out that at least sometimes the diagnosis of an error might be avoidable: mathematicians may assert certain counterpossibles like (Halting) and deny others like (Arithmetic) just in order to state their intuitions based on *prima facie* reasons, disregarding the issue of the possibility of the antecedent. I shall offer a further alternative in the next section (4.2.). All these alternatives taken together may well account for all cases in which mathematicians seem to commit themselves to non-vacuity.

(B2) It is not the case that an electron's angular momentum L in the ground state is observed at L=ħ.

(B3) Therefore, Bohr's theory of the atom is false.' (Tan [2019], p. 48)

This modus tollens argument resembles reductio ad absurdum arguments. (B1)–(B3) is perfectly sound if (B1) is vacuously true. This can be shown in an exemplary way using the standard possible world account: a counterfactual is true precisely if all closest antecedent worlds are consequent worlds. If the antecedent is actually true, the actual world is among the closest antecedent worlds. Yet the consequent of (B1) and (B2) cannot be true together. Hence, for (B1) & (B2) to be true, the actual world must not be a world in which Bohr's theory is true. Therefore (B1) & (B2) cannot be true unless (B3) is true. In sum, Tan has given no example of reasoning with superseded theories which requires non-vacuity.

One of the main targets of Tan's criticism is the claim that counterpossibles are uninformative. Indeed counterpossibles can be informative in the right context. Yet their information value is reconcilable with their vacuity. The informational value of (B1)–(B3) does not depend on the non-vacuity of (B1) but on choosing a concrete empirical prediction of Bohr's theory in the consequent of (B1) and in (B2). If Tan had used an uninformative statement like '2+2=5' instead, one would still infer that Bohr's theory is false, but one would not be informed about the empirical counterevidence that shows it to be false.

### 4.2 Counterpossibles read epistemically

My second counterstrategy supplements the first. As Tan notes, we often hypothetically reason from superseded theories in historical contexts:

'A physics instructor, for instance, might teach her students, [Stern–Gerlach] “If the classical mechanics of the atom had been true, then the Stern–Gerlach experiment would not have detected any abnormalities in angular momentum caused by spin.”

Since the instructor knows that the classical theory is false and aims to explain to her students why we know so, the utterance of this conditional’s antecedent is plausibly interpreted as her entertaining a counterfactual supposition.’ (Tan [2019], p. 50)

The quote leaves the instructor’s purposes underspecified. She might have in mind a modus tollens argument like (B1)–(B3). Then the considerations from the last section apply. The argument is informative not due to non-vacuity but because it describes the role of the Stern–Gerlach experiment. Another possibility of reading (Stern–Gerlach) without countenancing non-vacuity is the strategy developed in my discussion of Jenny. The physics instructor may use (Stern–Gerlach) simply to convey that there is a specific argument that leads her from the antecedent to the consequent without drawing on information on whether the antecedent is possible. In this case, the argument would lead from classical mechanics to its prediction the outcome of the experiment.

Still the following principled alternative is interesting: there is an increasing literature supporting that counterfactuals allow both for a metaphysical and an epistemic reading (Edgington [2011], Khoo [2015]). The former is characterized by the orthodoxy about counterfactuals. The latter tackles the antecedent as an epistemic possibility. It has even been argued that non-vacuity intuitions about counterpossibles can be explained by the epistemic reading (Vetter [2015], Williamson [2017], p. 217, criticism in Dohrn [2021]). I would not go so far, but if there is an epistemic reading, it provides a highly relevant alternative for reading some counterpossibles.[20]

Historical counterpossibles like the one about the Stern–Gerlach experiment may often be accounted for by an epistemic reading: in this reading, they say the same, albeit asserted at a later point in time, as corresponding indicative conditionals, uttered by someone in a past epistemic situation in which the truth of the pertinent theory was still a life option. The physics instructor in Tan's example may use her counterfactual to render the past epistemic viewpoint of a scientist to whom classical mechanics was still an option and the outcome of the Stern–Gerlach experiment (avant la lèttre) still open. That scientist might have said:

If classical mechanics is right about atoms, then the such-and-such (Stern–Gerlach) experiment will not detect any abnormalities in angular momentum caused by spin.

This conditional seems true and its denial false, and the same goes for (Stern–Gerlach), read in the way indicated.

To sum up my two new counterstrategies to Tan's argument from superseded theories: counterpossibles about false theories may be construed metaphysically; then nothing Tan says precludes their being informative albeit vacuous. Or they may be read epistemically; then the

[20] Williamson ([2020], 15.2–15.3) doubts that there is an epistemic reading of counterfactuals as distinguished from a metaphysical one. He suggests that the examples of such a reading can be interpreted by contextually restricting the possible worlds that are accessible for assessing a counterfactual to those which are also epistemically possible candidates for the actual world. However read, counterpossibles remain vacuous. If Williamson is right, my suggestion has to be dismissed. There are still the alternatives of doing with vacuous counterpossibles or with a *prima-facie* argument that does not draw on information whether the antecedent is possible.

question of metaphysical impossibility does not matter. We often reason with epistemic possibilities which turn out to be metaphysically impossible.

## 5 Conclusion

I have discussed the main extant arguments for the claim that scientific practice supports the non-vacuity of counterpossibles. I have used these arguments to illustrate several counterstrategies on behalf of non-vacuists. Although philosophy arguably should defer to scientific practice, it takes more than the examples adduced so far to establish that scientific practice supports the non-vacuity of counterpossibles.

## Acknowledgements

The paper has greatly profited from the constructive comments of two anonymous reviewers. This research was funded by the Department of Philosophy 'Piero Martinetti' of the University of Milan under the Project 'Departments of Excellence 2018–2022' awarded by the Ministry of Education, University and Research (MIUR).The author acknowledges support from the University of Milan through the APC initiative.

Daniel Dohrn

Dipartimento di Filosofia

Università degli Studi di Milano

Milan, Italy

Daniel.Dohrn@unimi.it

## References


Adams, E. [1993]: ‘On the Rightness of Certain Counterfactuals’, *Pacific Philosophical Quarterly*, **74**, pp. 1–10.

Baron, S., Colyvan, M., Ripley, D. [2017]: ‘How Mathematics can Make a Difference’, *Philosophers’ Imprint*, **17**, pp. 1–17.

Berto, F., French, R., Priest, G., Ripley, D. [2018]: ‘Williamson on Counterpossibles’, *Journal of Philosophical Logic*, **47**, pp. 693–713.

Bokulich, A. [2011]: ‘How Scientific Models Can Explain’, *Synthese*, **93**, pp. 33–45.

Davis, M. [1958]: *Computability and Unsolvability*, New York: McGraw–Hill.

Dohrn, D. [2021]: ‘Are Counterpossibles Epistemic?’, *Pacific Philosophical Quarterly*, **102**, pp. 51–72.

Edgington, D. [2011]: ‘Causation First: Why Causation is Prior to Counterfactuals’, in C. Hoerl, T. McCormack, and S.R. Beck (eds), *Understanding Counterfactuals, Understanding Causation: Issues in Philosophy and Psychology*, Oxford: Oxford University Press, pp. 230–41.

Godfrey-Smith, P. [2020]: ‘Models, Fictions, and Conditionals’, in A. Levy and P. Godfrey-Smith (eds), *The Scientific Imagination*, Oxford: Oxford University Press, pp. 154–77.

Goodman, N. [1947]: ‘The Problem of Counterfactual Conditionals’, *The Journal of Philosophy*, **44**, pp. 113–28.

Hesse, M. [1963]: *Models and Analogies in Science*, London: Sheed and Ward.

Jenkins, C.S. [2010]: ‘Concepts, Experience and Modal Knowledge’, *Philosophical Perspectives*, **24**, pp. 255–79.

Jenkins, C. S., Nolan, D. [2012]: ‘Disposition impossible’, *Noûs*, **46**, pp. 732–53.

Jenny, M. [2018]: ‘Counterpossibles in Science. The Case of Relative Computability’, *Noûs*, **52**, pp. 530–60.

Khoo, J. [2015]: ‘On Indicative and Subjunctive Conditionals’, *Philosophers’ Imprint*, **15**, pp. 1–40.

Lewis, D. [1973]: *Counterfactuals*, Oxford: Blackwell.

McLoone, B. [Forthcoming]: ‘Calculus and Counterpossibles in Science’, *Synthese*.

Nolan D. [1997]: ‘Impossible worlds: a modest approach’, *Notre Dame Journal of Formal Logic*, **38**, pp. 535–72.

Psillos, S. [2011]: ‘Living with the Abstract: Realism and Models’, *Synthese*, **93**, pp. 3–17.

Rogers, H. [1967]: *Theory of Recursive Functions and Effective Computability*, New York: McGraw-Hill.

Stalnaker, R. [1968]: ‘A Theory of Conditionals. Studies in Logical Theory’, *American Philosophical Quarterly Monograph*, **2**, pp. 98–112.

Stalnaker, R. [1981]: ‘A Defense of Conditional Excluded Middle,’ in W.L. Harper, G. Pearce, and R. Stalnaker (eds), *Ifs: Conditionals, Belief, Decision, Chance, and Time,* Dordrecht: Reidel, pp. 87–104.

Tan, P. [2019]: ‘Counterpossible Non-Vacuity in Scientific Practice’, *The Journal of Philosophy*, **116**, pp. 32–60.

Vetter, B. [2015]: ‘Counterpossibles (Not Only) for Dispositionalists’, *Philosophical Studies*, **173**, pp. 2681–700.

Williamson, T. [2007]: *The Philosophy of Philosophy*, Oxford: Blackwell.

Williamson, T. [2017]: ‘Counterpossibles in Semantics and Metaphysics’, *Argumenta*, **2**, pp. 195–226.

Williamson, T. [2017b]: ‘Model-building in philosophy’, in R. Blackford and D. Broderick (eds), *Philosophy’s future: The problem of philosophical progress*, Malden: Wiley, pp. 159–172.

Williamson, T. [2018]: ‘Counterpossibles’, *Topoi*, **37**, pp. 357–68.

Williamson, T. [2020]: *Suppose and Tell. The Semantics and Heuristics of Conditionals,* Oxford: Oxford University Press.

Williamson, T. [2020b]: 'Book Review—the Scientific Imagination: Philosophical and Psychological Perspectives', *Notre Dame Philosophical Reviews,* https://ndpr.nd.edu/reviews/the-scientific-imagination-philosophical-and-psychological-perspectives/

Wilson, A. [2018]: 'Grounding Entails Counterpossible Non-Triviality', *Philosophy and Phenomenological Research,* **96**, pp. 716–28.

Woodward J. [2003]: *Making Things Happen: A Theory of Causal Explanation*, Oxford: Oxford University Press.

Woodward, J. [2015]: 'Interventionism and Causal Exclusion', *Philosophy and Phenomenological Research*, **91**, pp. 303–47.

Yli-Vakkuri, J., Hawthorne, J. [2020]: 'The Necessity of Mathematics', *Noûs,* **54**, pp. 549–77.