

# The Contextual Character of Causal Evidence

Mauricio Suárez

Published online: 27 August 2013  
© Springer Science+Business Media Dordrecht 2013

**Abstract** I argue for the thesis that causal evidence is context-dependent. The same causal claim may be warranted by the same piece of evidence in one context but not another. I show this in particular for the type of causal evidence characteristic of the manipulability theory defended by Woodward (*Making things happen: a theory of causal explanation*. Oxford University Press, Oxford, 2003). My thesis, however, generalises to other theories—and at the end of the paper I outline the generalization to counterfactual theories. The paradigmatic form of causal evidence in the manipulability theory is provided by tests of the functional invariance of the relation between putative cause and effect under interventions (on the putative cause). I show that such evidence exhibits at least two kinds of context-relativity: personal, or epistemic; and situational, or objective.

**Keywords** Causal inference · Evidence · Contextualism · Interventions · Manipulability theory of causation · Jim Woodward

## 1 The Main Claim

Here are a few *prima facie* unproblematic causal claims: aspirins relieve headaches, bread nourishes, children raise anyone's spirits, inflation reduces purchasing power, forces applied to objects set them in motion. Most of us believe most of these claims, at least to some extent. Why? What are our grounds for such beliefs? Since the claims are empirical, our grounds must ultimately derive from

empirical evidence, even if we ourselves may have acquired some of those beliefs from the testimony of those we trust, as opposed to our own experience. In other words, whatever their provenance, these beliefs are justified by particular pieces of empirically accessible data. And since the content of those beliefs is causal, we may expect any theory of causation to discharge an obligation to enlighten us regarding the nature of such evidence. And indeed, most current theories of causation implicitly or explicitly propose their own account of the nature of causal evidence.<sup>1</sup>

In this paper I argue that causal evidence is contextual, i.e. it is relative to a context. More specifically, I defend the view that the same causal claim may be warranted by the same piece of empirical evidence in one context but not another. This view must be distinguished from the superficially similar view that causal claims themselves are contextual, or relative to context. My claim is about the context-relativity of evidence, not about the relativity of truth. I do not argue that what is a true causal claim in one context may be a false claim in another. On the contrary, I do not believe that the concept of truth in general is relative. That is, I do not believe that the truth conditions of declarative statements are generally relative to context—although I do of course accept that there are truths *about* particular contexts.

I also find that causal claims are amongst the least relative, uncertain, or indeterminate truths that we possess. We do indeed possess large amounts of causal knowledge about all kinds of particular and general systems and objects. And although such knowledge maybe very specific and singular, it does not seem to me to be relative in any particularly radical sense. More specifically, I do not believe that the examples that I provided above are relative

---

M. Suárez (✉)  
Department of Logic and Philosophy of Science, Complutense University, 28040 Madrid, Spain  
e-mail: msuarez@filos.ucm.es

<sup>1</sup> Throughout the paper “causal evidence” refers to whatever evidence may support a causal claim, whether token or type.

or contextual truths. Rather it seems to me that the same true proposition is expressed in every context by the statements that aspirins relieve headaches, bread nourishes, children raise anyone's spirits, inflation reduces purchasing power, and forces put into motion those objects that they act upon. There is no interesting sense in which these are socially constructed truths, or truths only for particular communities, practices or cultures.

None of the above entails that causal claims are always, or may always be understood as, binary claims involving a two-place relation between the two relata (the 'cause' and the 'effect', typically understood to be events). Many—maybe even all—causal claims are contrastive, in a sense that has been extensively discussed by, amongst others, Hitchcock (1995), Menzies (2007) and Schaffer (2005). Their contrastive nature is particularly conspicuous in some token causal claims such as “taking this aspirin relieves my headache”, or “force  $f$  impressed upon object of mass  $m$  imparts on it an acceleration  $a$ . The truth of these claims often depends on the contrast class implied. Thus “taking the aspirin rather than nothing” is one possible contrast class for the putative cause, while “taking the aspirin rather than the paracetamol” is another. Similarly that the aspirin will relieve “my headache rather than my stomach burn” provides one contrast for the putative effect, while “my headache rather than my sore throat” provides another. The original binary claim thus unfolds into several different contrastive claims, and the truth-value will sensitively depend upon the implied contrast classes in each and every case.

This is all true and plausible, yet notice that the contrastive nature of many, if not all, causal claims does not entail that their truth is relative to context. Rather it points out that many ordinary causal locutions involving binary relations are incomplete, in the sense that they are, in practice, elliptical for more complex locutions involving three or four place relations. More generally, the claim “ $c$  causes  $e$ ” is often elliptical for “ $c$  rather than  $c'$  causes  $e$  rather than  $e'$ ”, and this latter expression, which makes explicit the full causal claim intended, is objectively true or false, having definite and precise truth conditions and a truth value. These contrastive expressions are not relative to context in the way I argue in this paper evidence for them typically is. The contextualist thesis regarding causal evidence that I am defending in this paper is different, and independent, from the thesis that causal claims are contrastive in nature.

I will consequently assume here that our knowledge of causal truths is as absolute and context-independent as any other knowledge we possess, because we can only possess knowledge of what is in fact true in the non-relative or absolute sense described above. Yet, although my views about truth in general—and causal truth in particular—are conservative, my views about evidence—in particular causal evidence—are not. I do believe that evidence for

causal truths is contextual in an interestingly radical sense. Maybe this is also a sense in which any evidence for any claim may be said to be contextual; if so, it is still the case that the contextual nature of causal evidence has not been appreciated sufficiently so far.

My most general claim is that causal evidence is contextual *however* causal evidence is understood. In other words, I defend that the same causal claim may be warranted by the same piece of empirical evidence in one context but not another—and that this is so *independently* of the theory of causation, and causal evidence, that is adopted. There will always be important contextual presuppositions determining both what I call default entitlements and relevant confounding factors. In this paper I only have enough space to argue for the claim in connection with the manipulability theory of causation developed by Jim Woodward (2003). However, in the last section of the paper I go on to outline an extension of this argument to another theory, namely the counterfactual theory. This suggests that the context determines the objective standards required for evidence for a particular causal claim, regardless of how causal evidence is understood.<sup>2</sup>

## 2 The Manipulability Theory of Causation

In a rough form, Woodward's theory may be summarized in the following slogan: 'causation is invariance under intervention'. Given two variables  $a$  and  $b$  representing the putative cause and effect, we may say that  $a$  causes  $b$  if the functional relationship between  $a$  and  $b$  remains invariant under interventions on the putative cause  $a$ . I will discuss the various presuppositions of this theory later on. For now we may focus on the definition just provided, and proceed to clarify the terms of the theory in line with Woodward's own explanations.

The first thing to notice is that the slogan above is not a definition. Invariance under intervention is meant as a sufficient but not necessary condition for causation. There may be causal relations that do not entail such invariance. If causation does not entail invariance then testing for lack of invariance does not necessarily inform us regarding a lack of causation. Admittedly, Woodward is not entirely clear or explicit on this point. In earlier parts of the book he asserts that invariance is necessary: “If a relationship is to qualify as causal, it must be invariant under some interventions” (pp. 69–70). However, later on in the book this assertion is qualified by introducing “approximate

<sup>2</sup> I have always been inclined towards the view that epistemic justification is relative to a context, and Suárez (1999) already discussed and defended Michael Williams' (1991) brand of contextualism. But I did not there apply contextualism to specifically causal evidence.

invariance” (p. 239), and “degrees of invariance” (p. 257ff), which on the face of it turns exact invariance, or invariance per se into a non-necessary but merely regulative condition on causation. In the remainder of the paper I assume that invariance is not strictly a necessary condition on a causal relation, but that it is nonetheless sufficient. Such a weakening of the invariance condition does not affect what I say here regarding causal evidence.

Secondly, note that the causal relata on this theory are variables, which take values within a range. We may consider the variables to represent type-causes, with the taking of particular values representing singular or token causes. Thus ‘aspirins relieve headaches’ is a general type-level causal statement, while ‘this aspirin relieves my headache’ is a singular token-level causal statement. According to the manipulability theory only the former is represented by a generic “*a* causes *b*” with *a* standing for ‘taking aspirins’ and *b* standing for ‘relieving headaches’. Nevertheless there is a standard translation rule that takes us from the view that causal relata are events to the view that they are variables: every event can be represented in terms of a bi-valued variable that takes the value 1 when the event takes place and the value 0 when the event does not take place. The general type-level versus singular token-level views are then only distinguished in terms of the sorts of events that take place. In a general type-level statement of causation the events are event-types (as in “aspirin has been ingested somewhere”, “someone’s headache has been relieved”) while in a singular token-level statement the events are tokens (as in “I have just now ingested this aspirin”, “my headache is now gone”). Both can in principle be represented as variables taking values in such a straightforward trivial way (although see some complications later on), so I will not hereto distinguish between these two levels of causation, including in the relevant examples.

Finally, we need to define the concept of an intervention, and Woodward’s theory is particularly precise and helpful in this regard. Let us suppose that we are considering the putative causal relation between *a* and *b*. Then an intervention is a further variable defined by means of a number of conditions on its relation to both *a* and *b*. More specifically we say that the variable *I* is an intervention variable on *a* with respect to *b* if and only if:

- (i) *I* is a cause of *a*.
- (ii) *I* acts as a switch for any other variable *x* that causes *a* (within a certain range of values of *a*).
- (iii) Any directed path from *I* to *b* goes through *a*.
- (iv) *I* is statistically independent of any variable *x* that causes *b* and is on a directed path that does not go through *a*.

A few comments on these conditions are in order. First of all, (i) stipulates that *I* must act as a cause of *a*, which

means that no definition of causation in terms of invariance under intervention would constitute an analytical reduction of the concept, since it would include the term to be defined in the definiendum. It also means that no variable *v* can be an intervention variable if it fails to be causally related to the putative cause *a*, regardless of how highly *v* may be correlated with *a*. The fact that an intervention requires the actual existence of causal relations over and above any correlation has implications for the contextual character of causal evidence, as will be discussed later on.

As regards the remaining three conditions, I have elsewhere illustrated them graphically by means of diagrams representing their possible violations (Suárez forthcoming). Condition (ii) stipulates that no other cause of *a* may co-determine the value of *a* when *I* is acting; in other words, *I* is entirely responsible (a ‘total’ cause) for setting the value of *a*. This either rules out the existence of any other acting cause on *a* simultaneous with *I*, or it switches off the operation of such causes. Either way, it is required that as a matter of fact *I* actually fix the value of *a*. Condition (iii) also requires that some particular state of affairs obtains, namely: that the only directed paths from *I* to *b* all pass through *a*—this excludes any directed paths from *I* to *b* that bypass *a*. Finally condition (iv) stipulates that any independent cause of *b* must be statistically independent of the intervention variable *I*, thus ruling out in actual fact any statistical correlation between *I* and any cause of *b* other than *a*.

Let us apply the conditions to our main example. In the manipulability theory we say that aspirins relieve headaches if it is possible to intervene on aspirin-taking, by e.g. intervening upon a subject about to ingest an aspirin, in such a way that (i) the intervention causes the subject to ingest the aspirin, (ii) the intervention overrides any other cause that may prevent the individual from ingesting the aspirin, (iii) the intervention does not itself in any way relieve the individual’s headache, other than in virtue of causing the ingestion of the aspirin, and (iv) the intervention is statistically uncorrelated with any other factor that may help relieve the individual’s headache (such as exercising, or taking a walk in fresh air, etc.). The example shows that any unaccounted intermediate causal variable (i.e. any one that is not explicitly represented in the simple cause-effect model here described) between the ingestion of the aspirin and the relief of the pain, such as chemical processes occurring in the human body, would reveal itself under these conditions. It would show in a lack of invariance of the relation between the putative cause (aspirin) and putative effect (pain relief). Thus, conversely, a positive test of invariance under these conditions is prima facie evidence that the cause and effect variables are indeed causally related.

For our purposes here, it matters that these four conditions are all factual—they do not in any way refer to the

content of our knowledge, nor do they reflect any constraints or limits on our epistemic access to causal claims. They rather concern the states of affairs that must obtain in the actual world for  $I$  to be properly an intervention variable on  $a$  with respect to  $b$ . In other words, they describe the objective facts that must obtain for any tests of invariance (of the functional relation between  $a$  and  $b$  under interventions on  $a$ ) to constitute evidence for or against any causal connection. As I will argue in the next two sections, they provide us with the objective context presupposed for causal evidence.

### 3 Sources of Epistemic Contextuality

I argue that causal evidence is contextual. But what does it mean to say of any evidence that it is contextual? I borrow my understanding of contextualism from Michael Williams, who has for decades now employed it as part of an argument against a view in epistemology that he refers to as ‘epistemological realism’. So let me make it clear from the start that while I believe Williams has developed the best account of contextual evidence on offer, I do not necessarily endorse his epistemological conclusions against epistemological realism.

In *Problems of Knowledge* (2001), Williams argues that the sources of contextuality of evidence are twofold. There are what we may call personal and situational sources. Personal sources include those default entitlements required for particular agents to possess personal justification in their beliefs in their particular circumstances. They are part of the epistemic structure of the belief system of agents, and are not related to the features of the situation itself. To put it roughly, personal factors reside not in the world, but in the epistemic scaffolding of the cognitive processes of agents. Situational sources of contextuality are, by contrast, beyond the agent’s purview, reach, or control. They belong to the situation itself about which knowledge is claimed, not in any way to the knowing agent, his or her cognitive structure.

Thus, to continue with our example, suppose someone requests evidence for the claim that aspirins relieve headaches. For any piece of evidence to be effective, a set of default entitlements must be in place in the agents requesting and providing the evidence. The standards and type of evidence that will be appropriate will differ depending on these factors. The agents must both apprehend the referent of the terms in the claim. Thus they both must agree on the sort of entity that ‘aspirin’ refers to, including (at least a subset of) its properties. They must share an understanding of the sort of pain involved in headaches, and what would count as relief. They must also share an understanding of what sort of method is

appropriate in testing for the effect (i.e. what would be an appropriate method to figure out when in fact a headache has been relieved—whether it includes merely questioning of the patient, or more in-depth analysis of motor or neuronal responses or configurations, etc.—including, possibly, the case where the patient is one of the agents themselves). And they must be equally clear on the dialectics of their inquiry, since one is requesting the evidence and the other is providing it—what it would take to convincingly provide it, and to convincingly accept it as such, etc.). Finally, they must be clear on the risks and costs involved in the inquiry—in particular the risks and costs involved in relieving the particular headache in question. Medical cases make this particularly clear. Thus if a patient’s health condition is delicate there will be a premium on getting the diagnostic right, so the standards for evidence will rise. Whereas if the situation is an urgent one of life or death, the pressures on trying anything out quickly may be greater, so the standards required for evidence may correspondingly decrease.

Williams thus distinguishes four different types of personal sources of contextuality: (i) intelligibility, or semantic; (ii) methodological or concerned with the rules of the particular inquiry; (iii) dialectical, i.e. related to the particular features of the specific space of giving-and-taking of reasons involved in the exchange; and (iv) economical, or pertaining to the cost-benefit analysis or evaluation of risks. He uses a simple example to illustrate them all, which is worth quoting in full as follows (Williams 2001, pp. 162–164):

Seeing someone drive by an old sports car, two people engage in the following exchange:

A: Isn’t that old sports car an E-Type?

B: Yes, a rare early model.

A: What makes you say that: don’t they all look pretty much the same?

B: Sure, but that one had external bonnet latches which you only get on the first five hundred cars.

Here A and B concede to each other various default entitlements—that an old car just passed by, and so on. Without such concessions, their conversation could not take the specific direction it does take (methodological factors). Indeed, if certain things (e.g. the capacity to tell a car from an elephant) could not be taken for granted, the speakers could not have any kind of intelligible exchange about the vehicle passing by (semantic factors). But B makes a claim implying special expertise [...] If, as A suspects, B’s remark was prompted by a casual glance at the passing car, this is an epistemic defeater. Citing it challenges B to back up his claim [(dialectical factor)]. B’s response is sufficient to meet A’s challenge.

[...] But if A were thinking of buying a car (an economic factor), this might [need revision in light of further factors].

The example brings into relief all the personal factors that need to be in place for the observation of particular properties of the vehicle to count as evidence that it is of a particular type. The point Williams is making here is that only the context offers the resources to understand why such observation is rightly appreciated as convincing personal justification for the belief in the claim in question. For, in a different context, where different default entitlements are in place, the same observation would not provide personal justification. It would not provide justification if the semantic factors were lacking that allow the agent to understand the meaning of the terms involved in the claim; or the methodological factors that allow the inference from the observed features to a certain distinct production; or the dialectical factors that make such an observation relevant at that point in the inquiry; or the economic factors that make the outcome of the inquiry salient to the agent's interests. Were those factors different, justification may well be lacking, or uncalled for. Or, a different sort of observation or procedure would be required for personal justification. The point Williams is therefore making is that the character of evidence itself is contextual, as in the example above—whether an operation or observation counts as evidence for a particular claim depends on the features of the context.

The point extends to objective sources of evidence, as Williams himself explains. For any claim regarding a particular state of affairs, there are objective interfering or confounding factors in the situation, in addition to those personal factors we reviewed, such that, were they present, it would be impossible for agents to provide and acquire evidence in favour, or against, that particular claim. Evidence requires not just standards of personal justification to be in place, it also requires a degree of reliability in the procedures that bring about the evidence and connect it with the state of affairs about which knowledge is claimed. These procedures provide a part of the contextual background of the evidence too.<sup>3</sup>

In Williams' own example, possible confounding factors would include the existence of a replica of the car in question, or the additional production, unknown to the

agents, of a further series of cars with the prescribed characteristics (external bonnet latches). More generally, confounding factors may be identified with those factual matters that affect the production and acquisition of reliable evidence. Default assumptions will be made regarding the absence of such factors, but they typically cannot be ruled out entirely by any additional evidence, so they will tend to be presupposed. As we shall shortly see, causal evidence is in a very straightforward way contextual in this objective, or situational, sense. The manipulability theory, in particular, provides a comprehensive list of the types of confounding factors that must be ruled out in advance for causal evidence to be possible.

#### 4 How Causal Evidence is Contextual According to the Manipulability Theory

In this section I show that causal evidence in the manipulability theory is contextual in at least two different ways: personal and situational.<sup>4</sup> Amongst the personal sources I focus only on methodological factors, since only these are characteristic of the manipulability theory. (The remaining personal factors are either too general—as in the economic factors, which seem to go beyond any of the manipulability theory's tenets—or more specific—in the case of the dialectical factors specific to the giving and taking of justificatory reasons amongst the agents). I then discuss the situational objective factors that underpin the contextual nature of causal evidence according to this theory. The combination of them all provide a rich tapestry of the default entitlements and excluded confounding factors that must be presupposed for causal evidence.

##### 4.1 Default Methodological Entitlements

The methodological entitlements of the manipulability theory are those premises that enable the application of the theory to any putative causal claim. The following is not an exhaustive list but includes what I take to be four basic methodological tenets, which provide the context of application of the theory (in each case I have emphasised the most relevant terms):

<sup>3</sup> Every (respectable) brand of contextualism accepts that there are situational or external contextual elements required for knowledge, going well beyond the merely personal factors previously discussed. Contextualism is not at all relativism about knowledge. For example, on Lewis' account (Lewis 1996, p. 554), S knows that P only if the evidence eliminates every possibility in which not-P—except those possibilities that are properly ignored in the context. Yet, Lewis' *rule of actuality* stipulates that no possibility that obtains as a matter of fact can be properly ignored. Hence evidence is always relative to the actual facts in the context—for the facts cannot be discounted.

<sup>4</sup> I believe that there is in fact a distinct third source of contextuality, which is related to the interpretation of a theory, and I refer to as hermeneutical, in those cases in which causal relations are predicted or described by an explanatory theory. The quantum mechanical correlations provide a striking example, since they may be described in very different terms in either orthodox quantum mechanics, or a variety of interpretations, including Bohmian mechanics. But this additional source of contextuality only adds to my main thesis, and need not detain us here.

1. Causes and effect are *variables* in a directed acyclic graph.
2. Variables take *values* within particular *ranges* (significant *intervals* for the assessment of invariance).
3. Variables may have *quantitative functional* relations over such significant intervals.
4. Interventions are in principle *possible* upon (the values of) putative causes.

Earlier in the paper I mentioned that any theory of causal relata as events can in principle be turned into a theory of variables, by adopting bi-valued variables. However, the translation is not always faithful, in that some meaning may be lost in the process. Often what matters in a causal claim is not the occurrence of an event simpliciter (e.g. the taking of an aspirin simpliciter), but the occurrence of an event-rather-than-another (the taking of aspirin as opposed to paracetamol, or vitamin; or as opposed to not taking anything). In such cases the relevant contrast class may not always be described by means of variables without loss of significant causal content.

Some cases of contrastive causation certainly have been treated by means of variables in the literature. Peter Menzies (2007, p. 208), for instance, proposes a general account of difference-making in terms of the structural equations formalism. Amongst his examples, at least two (examples “C” and “D” on his list) are explicitly contrastive and accounted for by means of 3-valued variables. Case C is Hesslow’s much-discussed example concerning contraceptive pills as a cause of thrombosis, and a three-valued variable C is introduced to account for the difference between ‘taking contraceptive pills-rather-than no contraception’ versus ‘taking the pill rather-than using some other contraceptive means’. While I do not doubt that many—and perhaps even ultimately all—causal claims are contrastive, I do doubt that all such claims can be treated exhaustively by means of multi-valued variables. After all, the contrast class is likely infinite or indefinite in many relevant cases; and many of those other relevant cases that possess definite and finite contrast classes can be treated as variables only trivially, by means of on–off variables.

At any rate, the discussion of the contrastive nature of causation is a tangential issue, since I only use it in order to illustrate the following general thought. The assertion that the relata of causation are variables is a genuine presupposition of inquiry, which not every causal context need to satisfy—or at least not in any straightforward way. Since it is part of the context for the application of the manipulability theory, it therefore also implicitly fixes the context for the application of the notion of causal evidence in question. The context must be such that the assertion is satisfied—i.e. it must be such that it makes sense to represent the causal relata as variables. Otherwise the theory is

not applicable and causal evidence cannot take the form of invariance under interventions. While this presupposition may be common, and conspicuously useful for analysis, it nonetheless remains a presupposition.

Similarly for the other three items in the list. Each establishes a methodological posit of the manipulability theory. Together (2) and (3) entail that  $I, a, b$  are all variables taking values within particular intervals over which some functional relation  $F(a, b)$  ranges between  $a$  and  $b$ . This may be false even in contexts that allow us to represent the relata as variables, as required by (1). For instance, many contexts in which the natural way of thinking about relata is in terms of events, will admit a simplified translation into some variables  $a$  and  $b$  (as discussed above) but not to those variables that range over intervals such that some functional dependency between  $a$  and  $b$  holds. The example at hand is arguably one such: ‘taking aspirin’ and ‘relieving headaches’ may be represented as bi-valued variables, but not variables ranging over intervals where a functional relation of the sort  $b = ma$  (with  $m$  some constant) will typically hold.<sup>5</sup>

As regards (4) it is a well-known feature of the manipulability theory that interventions need not have been carried out but must merely be possible (see e.g. Woodward 2003, p. 127). There are intriguing and subtle issues surrounding the notion of possibility that is at play at this point (Waters 2007, p. 560ff; Samaniego 2011, Ch. 4). However, for our purposes here we can take those debates as settled (or irrelevant) by adopting the most restrictive sense of ‘possible’—i.e. as possible within the actual world with the kind of technological means that we now know ourselves to possess. The reason is that we are concerned in this paper not with the truth conditions of causal claims but with the conditions for causal evidence. What is actually true may well depend on what is possible in a robust modal sense—at least it will do if there are modal truths. But what is actually warranted cannot in the same way depend on what is possibly warranted in this sense. If we are to have *actual* evidence of causation, we must have to hand actual tests of invariance. And this requires *actual* interventions, so the availability of evidence minimally presupposes that interventions are possible in the restricted sense above. May be not now, maybe not here—but if we, as knowing agents, are to possess any evidence that  $a$  causes  $b$  then we must (under the appropriate circumstances) be in fact able to actually intervene on  $a$  in order to test the invariance of the functional connection  $F(a, b)$  between  $a$  and  $b$ .<sup>6</sup>

<sup>5</sup> The problem here does not lie with indeterminism, for even a probabilistic relation, such as  $Prob(b) = m Prob(a)$ , seems inappropriate as a representation of this particular functional relation.

<sup>6</sup> The careful formulation above leaves it an open question whether we have in fact performed the interventions. For instance, we may not have performed them on account of not having found ourselves in the

## 4.2 Default Situational Factors

Let me now address the objective or situational confounding factors that must be in place for the possibility of causal evidence. I already mentioned that some facts about the situation can not be discounted, but that evidence must be defined relative to them. One of the nice things about the manipulability theory is that it provides a ready-made list of those contextual factual factors that may not be ignored. For as we saw there are four factual conditions that must be satisfied for a variable  $I$  in a causal graph to count as an intervention. These four conditions provide us precisely with a guide to the array of arrangements of the variables in the graph that would prevent us from gaining any evidence by means of invariance tests. In other words: if any of these four conditions is not met as a matter of fact then no positive test for the invariance of  $F(a, b)$  while appropriately fixing the values of the ‘intervention’ variable can qualify as evidence for the causal claim regardless of whether all the personal contextual factors are in place. Let me remind the reader what these conditions are, in the contrapositive form, which serves to best describe them as part of the contextual presuppositions of evidence.

The variable  $I$  is not in fact an intervention on  $a$  with respect to  $b$  (and thus the fixing of its value is an ineffective way of gaining evidence of a causal connection between  $a$  and  $b$ ) if either of the following conditions obtains:

- (i)  $I$  is not in fact a cause of  $a$ .
- (ii)  $I$  does not switch off all other causes of  $a$  and, as a consequence, it does not in fact fix the value of  $a$ .
- (iii)  $I$  is an independent cause of  $b$  and hence in fact fixes the value of  $b$  by means other than its fixing the value of  $a$ .

Footnote 6 continued

appropriate circumstances. Nonetheless the interventions must be possible in the restricted sense above—which amounts to not more than a sense of practical *situational* possibility. Hence the sort of *hypothetical* experiments invoked in Woodward (2003, e.g. on p. 35) as providing causal evidence, are understood here to be hypothetical in this restrictive practical sense—they are experiments that could be carried out now, in this world, with our present technological capabilities. They may only be said to be counterfactual in this very limited modal sense. More specifically, the only actual evidence of causation we may possess that goes beyond those tests of invariance performed by us under our own interventions, is provided by analogical reasoning from tests that have been naturally performed elsewhere by naturally occurring interventions. This allows for causal evidence in fields like astrophysics, for example, wherever there are convincing reasons to believe that naturally occurring interventions are taking place. (Gravitational lensing is a good example of a naturally occurring intervention of this sort). The restrictive sense of possibility adopted here is enough to account for the cases of possible interventions that, it has been suggested, are sources of evidence for causal claims. And it fortunately does not require any problematic appeal to possible world semantics—i.e. no appeal is needed to interventions that only occur in worlds other than the actual world.

- (iv)  $I$  is correlated with an independent cause  $x$  of  $b$ , and hence is correlated with the value of  $b$  independently of it fixing the value of  $a$  and  $a$  being in fact a cause of  $b$ .

In any of these cases, the functional relation  $F(a, b)$  may remain invariant under operations to change the value of  $a$  by means of  $I$  within the prescribed ranges without  $a$  in fact being a cause of  $b$ . Therefore invariance under intervention in any such case does not constitute evidence in favour of a causal connection *regardless* of any other personal epistemic contextual factors—and in particular regardless of whether all the methodological entitlements of the manipulability theory are in fact met.

Let me then return to the simple example for illustration. Suppose that we have resolved all issues regarding the methodological application of the manipulability theory to the example. We then proceed to run repeated experiments on patients with severe headaches. And suppose that in these experiments we entirely and reliably control for the patients’ ingestion of aspirin by means of some exogenous variable, such as e.g. dissolving it in their drink. Now suppose that for those patients this variable is in turn a cause of relieving their headache (which we may suppose is due to dehydration and therefore relieved by the mere ingestion of liquid). We can easily check for the deterministic relation between aspirin ingestion and headache relief in these conditions—and we find that the relation is invariant: ingestion is invariably followed by the pain relief. This is indeed the sort of invariance that would typically constitute evidence of a causal relation on the manipulability theory. Yet, in this particular context, and on the contextualist view defended here, even though it is true that aspirins have the power to relieve headache, the observation of invariance across all the patients does not in fact constitute evidence for the causal claim. The ‘intervention’ variable is not in fact an intervention, since it does not satisfy condition (iii) upon interventions. Instead it causes the putative effect (pain relief) independently of the putative cause (aspirin ingestion). For all these patients the mere ingestion of the drink without any aspirin would have resulted in the pain relief anyway.

Compare this situation to another one—identical in all other respects—where dehydration is not in fact causally relevant to the headache. In this alternative situation, the default situational factors change, and the mere ingestion of liquid cannot relieve headaches. The intervention variable is no longer an independent cause of the effect via a third route. On the contextualist view defended here, the invariance of the aspirin-taking and headache-relieving variables does count as a source of evidence for the causal claim in this situation. The very same observations upon the very same operations give rise to causal evidence in one

context but not another—and the difference between both contexts is independent of any personal epistemic factors, and down entirely to particular factual matters in the situation that distinguishes them. These facts regarding the underlying causal structure determine to some degree what observable features count as evidence for or against what claims.

## 5 Extension of the Argument

The example above illustrates how causal pre-emption can vitiating causal evidence in the manipulability theory. Similar cases will apply in other theories, and this can be shown easily by transferring the argument to those theories via some of the features they share with the manipulability theory. For instance, on a counterfactual theory (Lewis 1973) causation is analysed in terms of counterfactual dependence. According to that theory ‘*a* causes *b*’ is true if and only if: (i) *a* and *b* both take place, and (ii) had *a* not taken place then *b* would have not taken place.<sup>7</sup> Now, there are significant differences between Lewis’ counterfactual theory and the manipulability account (some are rehearsed in Woodward 2003, pp. 133–145). For instance, the counterfactual theory is meant as a reductive analysis of the concept of causation, which the manipulability theory is not. The relation between manipulability and counterfactuals is rather a one-way entailment: manipulability conditions will typically entail some counterfactuals, but not necessarily the same counterfactuals that entail them. Thus to say that the functional relation between *a* and *b* is invariant under interventions on *a* entails that were the value of *a* to be changed by intervening on it, the value of *b* would change in accordance to that very functional relation. But there is no similar counterfactual that would entail the manipulability condition, so a reduction (an equivalence entailment both ways) is simply impossible.

Yet, the one-way connection described above is sufficient to transfer the argument for the contextuality of causal evidence to the counterfactual theory. The transfer strategy goes roughly as follows. According to the manipulability theory, *a* causes *b* if, were an intervention performed to fix the value of *a*, the value of *b* would co-vary accordingly. But, as was noted, such a co-variation

may be taken as evidence in favour of a causal connection between *a* and *b* only if all the relevant default epistemic entitlements are in place, and none of the confounding factors obtain in fact. The list of default entitlements is by and large different in the case of the counterfactual theory—since it is not required that the causal relata are variables, that they take any particular range of values, or that there is any functional relation between them.<sup>8</sup> However, there is an equivalent requirement to the possibility of interventions, which comes to the fore as part of Lewis’ semantics for counterfactuals.

As is well known Lewis analyses counterfactuals by means of similarity relations amongst possible worlds. In figuring out whether *b* counterfactually depends on *a*, we must look at the closest, i.e. the most similar, possible world in which *a* fails to occur. If *b* also fails to occur in that world, then the counterfactual is true (in this world). There is a recommended recipe for assessing closeness, or similarity of worlds, which stipulates that worlds where there are violations of our actual world’s laws, or mismatches in whole spacetime regions, are less close or similar to our actual world than those in which there are only local differences in particular facts. Thus the most similar world for our purposes, is the world in which nothing differs from the actual world until the time of *b*’s occurrence—other than the occurrence of *a* itself. The introduction of such a particular difference is what Lewis calls a local ‘miracle’. But to say that a local miracle has taken place in that world is just to say that an intervention could have been made in this world to fix the value of *a*.<sup>9</sup>

The possibility of interventions is therefore a contextual default entitlement in the counterfactual theory, just as much as it is in the manipulability theory. Were such an entitlement to be excluded from the stock of presuppositions then no evidence would be forthcoming for the causal claim in question. For if an intervention is not possible then it is not possible to acquire any evidence about what would be most likely the case in the most similar possible world in which the antecedent of the counterfactual is in fact true. And once this entitlement is in place, it is easy to see that the default confounding factors that must be ruled out for interventions on the manipulability theory would ipso facto have to be ruled out too on the counterfactual theory. For instance suppose that there is no exogenous causal variable *I* that may act to determine the non-occurrence of *a*, either because the variable *I* is not in fact a cause (thus flunking requirement (i) on interventions above), or because it does

<sup>7</sup> The theory is slightly more sophisticated in reducing causation to the ancestral of counterfactual dependence. So *a* causes *b* if and only if (i) *a* and *b* both occur, and (ii) *a* is related to *b* by means of some causal chain {*a*, *c*<sub>1</sub>, *c*<sub>2</sub>, ..., *c*<sub>*n*</sub>, *b*} where each member in the chain is counterfactually dependent on the previous member. This definition is equivalent to the simplified one in the text only under the substantial assumption of the transitivity of the relation of counterfactual dependence. I ignore the complication since it is irrelevant to the point I make in this paper regarding causal evidence.

<sup>8</sup> There is of course the requirement that cause and effect be counterfactually dependent in the prescribed way. But counterfactual dependence is not a functional relation, or at any rate, not a quantitative one.

<sup>9</sup> As Woodward himself notes in (Woodward 2003, p. 135).



not ‘switch off’ other causes of *a* (as in requirement (ii) above). If so, then there is no miracle that can be performed to prevent *a*, and no evidence can be forthcoming regarding the likelihood of *b*. Similarly, if the miracle that prevents *a* in some possible world turns out to also independently prevent *b*, then a closer possible world cannot be ruled out in which a different miracle just prevents *a* without necessarily preventing *b*. And, acting in this way to prevent *a* would no longer provide any evidence for the causal claim in question, etc.

To sum up, if we are to gain any evidence for what holds in the closest possible world that makes a counterfactual conditional true, we must intervene so as to perform changes in the conditions that appear in the antecedent. But any evidence that we can derive in this way will then be contextually dependent on both default epistemic entitlements, and the absence of confounding factors—precisely as in the manipulability theory.

Since much hinges here upon the interpretation of the nature of evidence in those cases where, on the manipulability theory, no interventions are possible, a few words may be in order regarding Woodward’s own views about such cases.<sup>10</sup> As it turns out Woodward’s views on this matter are not unambiguous. I have dealt with some of these ambiguities in relation with the EPR correlations in quantum mechanics elsewhere (Suárez 2007, forthcoming). But, more generally, Woodward seems to oscillate between two different accounts of the causal facts in cases where no interventions upon some putative cause *C* with respect to a putative effect *E* are available, namely: (i) that in such cases no claim of the form “*C* causes *E*” makes sense (i.e. any such claim is meaningless), and (ii): that in such cases no claim of that form is true (i.e. any claim to the effect that *C* causes *E* is false).<sup>11</sup>

Both alternatives seem counterintuitive in ordinary cases, where we typically can easily make sense of the claims, and they do not look *prima facie* necessarily false. But what is worse, neither is applicable to those cases, like the one discussed above concerning aspirins, dehydration and headache relief, where *ex-hypothesis* a causal relation is postulated to exist between *C* and *E*, but no intervention variable satisfies the four requirements (i–iv) above. It seems to me that the present contextualist proposal regarding evidence provides a much better response to such

<sup>10</sup> I thank a referee for helpfully stressing the need to make explicit the comparison with Woodward’s views.

<sup>11</sup> For instance, in Woodward (2003, p. 128), it is stated that: “[...] If we cannot think of *X* as a variable that is capable of being changed from one value to a different value [...] then claims about what will happen to *Y* under interventions on *X* will either be false or will lack a clear meaning”. See also the discussion of causality in the physics of the entire universe in Woodward (2009, section 10), which considers whether an earlier state of the universe may be meaningfully said to cause a later state, given that no interventions on the whole universe are in any way physically possible.

cases. We can say that in those cases, *ex-hypothesis*, it is true that *C* causes *E*, but that there is no evidence available to us that would enable us to discover, discern or establish such claim, since no evidence is in principle available to us that would allow us to distinguish such cases from those where the intervention variables represent genuine interventions. Hence on this view, it is not the causal claim itself that is false, but rather the claim that invariance under intervention provides evidence for or against the causal claim. And neither the causal claim “*C* causes *E*” nor the evidential claim “the invariance of *C* and *E* under interventions *I* on *C* is evidence that ‘*C* causes *E*’” are in any way meaningless or indeterminate. Their being meaningless would be in line with Woodward’s pronouncements but, as pointed out, would both violate our intuitions, and the set-ups of some thought experiments. On the contrary, causal claims have, on the view defended here, absolute and non-contextual truth conditions and values at all times. Yet, there may be no way for an agent to determine their truth value on account of evidence—at least no way that is independent of both personal and situational factors within the agent’s particular context.

## 6 Conclusions

In this paper I have argued that, whereas the truth of causal claims is not necessarily relative to any context, the evidence for or against that causal claim is heavily dependent on the context of inquiry. The same observations, under the same interpretation, resulting from the same operations, performed in identical manner, will constitute evidence for a causal claim in one context but not in another.

I have shown this to be the case if causation is understood along the lines of the manipulability theory. In so doing, I have separated epistemic default entitlements from the objective confounding factors in the situation that must be ruled out in advance. The sources of the contextuality of causal evidence are thus seen to be at least two-fold. However, I claim that the thesis of the contextuality of causal evidence generalises, and that causal evidence is contextual even if causation was understood differently. I have gone on to provide the outline of an extension of the argument to another theory—namely the counterfactual theory of causation.

**Acknowledgments** I thank the audience at the Conference on Evidence and Causality in the Sciences at the University of Kent, in September 2012, for their reactions; Federica Russo and Phyllis Illari for their editorial comments and suggestions; and two anonymous referees for their comments. The paper was written while visiting the Institute of Philosophy, London University. Financial support is acknowledged from the Spanish Ministry of Science and Innovation project FFI-2011-29834-C03-01.

## References

- Hitchcock C (1995) The role of contrast in causal and explanatory claims. *Synthese* 107:395–419
- Lewis D (1973) “Causation” reprinted in D Lewis (1986) philosophical papers, 2nd edn. Oxford University Press, Oxford, pp 159–213
- Lewis D (1996) Elusive knowledge. *Aust J Philos* 74(4):549–567
- Menzies P (2007) Causation in context. In: Price H, Corry R (eds) *Causation, physics and the constitution of reality*. Oxford University Press, Oxford, pp 191–223
- Samaniego F (2011) Explanatory depth and statistical mechanical interventionism. Ph.D. Thesis, Complutense University of Madrid
- Schaffer J (2005) Contrastive causation. *Philos Rev* 114(3):297–328
- Suárez M (1999) Epistemology in the face of strong sociology of knowledge. *Hist Human Sci* 12(4):41–48
- Suárez M (2007) Causal inference in quantum mechanics: A reassessment. In Russo F, Williamson J (eds) *Causality and probability in the sciences*. London College Texts, London, pp 65–106
- Suárez M (Forthcoming) Interventions and causality in quantum mechanics. *Erkenntnis*
- Waters K (2007) Causes that make a difference. *J Philos CIV*(11): 551–579
- Williams M (1991) *Unnatural doubts*. Blackwell, Oxford
- Williams M (2001) *Problems of knowledge*. Oxford University Press, Oxford
- Woodward J (2003) *Making things happen: a theory of causal explanation*. Oxford University Press, Oxford
- Woodward J (2009) Agency and interventionist theories. In: Beebe et al (eds) *The Oxford handbook of causation*. Oxford University Press, Oxford, pp 234–264