

## Hume's Two Definitions: The Procedural Interpretation\*

**Please do not cite this version.** The published version is: 'Hume's Two Definitions: The Procedural Interpretation', *Hume Studies*, 37 (2011): 243-74

### 1. Introduction

Hume's two definitions of causation have caused an extraordinary amount of controversy. The starting point for the controversy is the fact, well known to most philosophy undergraduates, that the two definitions aren't even extensionally equivalent, let alone semantically equivalent. So how can they both be *definitions*? One response to this problem has been to argue that Hume only intends the first as a genuine definition – an interpretation that delivers a straightforward regularity interpretation of Hume on causation. By many commentators' lights, however, this is a bug rather than a feature: such an account of the two definitions leaves necessary connection out of Hume's story about the meaning of "cause" all together, and this hardly does justice to his contention that the idea of necessity is an essential part of the idea of causation. In the light of this problem, some commentators have claimed that Hume's intention is not to *define* causation at all. Rather, the "definitions" are characterisations of the *circumstances* in which causal judgements are made – the first definition characterising the external circumstances (priority, contiguity and constant conjunction) and the second characterising the internal mechanism – the "determination of the mind" – that those external circumstances trigger.

---

\* Many thanks to seminar audiences at Cork, Sheffield, Cardiff, Otago and ANU who heard and commented on earlier versions of this paper, and to the editors and two referees for this journal for very constructive and helpful criticisms and suggestions.

These and other interpretations of the two definitions have their problems. The problem I am interested in here, however – and it is a problem that has received very little attention in the literature – is that none of these interpretations do justice to Hume’s claim in the *Treatise* that the first definition characterises causation as a philosophical relation, and the second as a natural relation. I shall argue that if we take this claim seriously, then – once we have an adequate understanding of the natural/philosophical distinction that Hume is referring to – it is entirely transparent why Hume should have provided two “definitions”; and it is also entirely transparent why they do not need to be extensionally equivalent (or rather, the extensional equivalence issue does not so much as arise). The two definitions reflect two different mental *procedures* by which causal judgements come to be made, namely the comparison (first definition) and association (second definition) of ideas. I dub this account the “Procedural Interpretation”.

In his book, *Hume*, Terence Penelhum says, “In the *Treatise* ... there is an obfuscatory complication. Hume says that the first definition is of causation considered as a *philosophical relation*, and the second is of causation considered as a *natural relation*. This distinction makes no appearance in the *Enquiry*, and its absence is a benefit. But its presence in the *Treatise* is a substantial barrier to the understanding of that work.”<sup>1</sup> Whatever the benefits may be of the absence of the distinction in the *Enquiry*, a lack of obfuscatory complication is not one of them, or so I shall argue, because there is no obfuscation in the *Treatise*. Hume’s two definitions in the *Treatise* make perfectly good sense when understood in the light of the distinction between natural and philosophical relations.

I shall proceed as follows. I begin, in §2, by providing an account of Hume’s conception of the distinction between natural and philosophical relations, which conceives of the distinction as a psychological distinction between the comparison and the association of ideas. I contrast this with the standard view, which I criticise. In §3, I propose the Procedural Interpretation of Hume’s definitions as they appear in the *Treatise*, and show how it does justice to the distinction between natural and philosophical relations. I also point out some of the consequences of the interpretation, and respond to a possible objection. In §4, I argue that none of the most common interpretations of Hume’s two definitions can make good sense of his claim that they characterise causation separately as a philosophical relation and as a natural relation. Finally, in §5, I briefly say something about some potential concerns about the application of the Procedural Interpretation to the first *Enquiry*.

## **2. Hume’s theory of natural and philosophical relations**

Hume presents the distinction between natural and philosophical relations right at the beginning of the *Treatise*. There are, he tells us, seven philosophical relations (T 1.1.5; SBN 13-15<sup>2</sup>): resemblance, identity, spatial and temporal relations, quantity or number, degree of sameness of quality, contrariety, and causation. Of these, two – resemblance and causation – are also natural relations, and so is one specific kind of spatio-temporal relation, namely contiguity (T 1.1.4; SBN 10-13).

The natural relations, Hume tells us, are those “qualities, from which [the association of ideas] arises, and by which the mind is after this manner convey’d from one idea to another” (T 1.1.4.1; SBN 11). Thus, for example, if you’re looking at a picture of Bertrand Russell, your mind will naturally be drawn to thinking about

Bertrand Russell himself, thanks to the relation of resemblance. If you're thinking about last week's conference dinner, your mind might naturally be drawn to the person who was sitting next to you or to what happened afterwards; that's contiguity. And if you're thinking about stubbing your toe, your mind will be drawn to the idea of pain: causation. Resemblance, contiguity and causation, then, are "principles of association". They correspond to what might be called "associative mechanisms", such that when I believe or judge that one thing, *a*, stands in one of these three relations to something else, *b*, my having the idea of *a* before the mind will, thanks to the relevant mechanism, draw the idea of *b* to mind.

Philosophical relations, by contrast, are "any particular subject of comparison, without a connecting principle" (T 1.1.5.1; SBN 13). Purely philosophical relations (that is, those that are not also natural relations) do not, according to Hume, give rise to any such associations. If you're thinking about Paris, your mind is not naturally drawn to the idea of a city *bigger than* or *smaller than* Paris. If you see a red post box, your mind isn't naturally drawn to the idea of something that is a darker shade of red. And so on. However, we can *consider*, and form judgements about, philosophical relations. That is to say, I can consider whether Paris is bigger than Rome, I can believe that the post box is a darker red than my gloves, and so on; and my doing so is a matter of "comparing" ideas under the relation in question. And of course I can do this for those philosophical relations that are also natural relations: I can believe that we'll go to the pub after the seminar, or imagine getting a nasty stomach ache as a result of consuming too many strawberries, or wonder whether a picture of Julius Caesar really looks anything like him.

The interpretation of the distinction between philosophical and natural relations that I shall briefly defend in this section is one according to which the distinction is a *purely psychological* distinction. To describe a relation as a “philosophical” relation is to consider it as a “subject of comparison”, where comparison is a certain sort of mental activity; and to describe a relation as a “natural” relation is to consider it as a very different sort of mental activity, whereby one idea naturally introduces another. Or, as Hume himself puts it:

The word *relation* is commonly us'd in two senses considerably different from each other. Either for that quality, by which two ideas are connected together in the imagination, and the one naturally introduces the other ... or for that particular circumstance, in which, even upon the arbitrary union of two ideas in the fancy, we may think proper to compare them. (T 1.1.5.1; SBN 13)

Here, Hume clearly distinguishes between natural and philosophical relations on the basis of two distinctive kinds of mental operation: on the one hand, a mechanism by which one idea naturally introduces another, and, on the other, the particular circumstance in which *we* “think proper to compare” two ideas.

This conception of the natural/philosophical distinction as a purely psychological distinction will become clearer by explaining how it differs from standard interpretations. One standard interpretation gains textual support from the sentence immediately following the passage quoted above:

In common language the former is always the sense, in which we use the word, *relation*; and 'tis only in philosophy, that we extend it to mean any particular subject of comparison, without a connecting principle. (T 1.1.5.1; SBN 13)

This sentence has been taken to suggest that Hume intends the natural/philosophical distinction to mark a distinction in ordinary language between what we normally think of as a “relation” and what the philosopher, who has a different agenda, thinks of as a relation. Thus Jonathan Bennett says that for Hume, “any true statement of the form  $xRy$  ... asserts the holding of a philosophical relation between  $x$  and  $y$ ”, whereas “there is a natural relation between  $x$  and  $y$  only if the plain man would describe them as ‘related’ or ‘connected’ ”.<sup>3</sup> In similar vein, Penelhum says, “Hume clearly supposes that philosophical relations are relations we speak of because we find them to be present on special examination, whereas plain men do not mention them because they do not notice them”.<sup>4</sup>

Such an interpretation unfortunately fits extremely uneasily with the use to which Hume puts the natural/philosophical distinction. Consider the passage that concludes Hume’s famous discussion of causal reasoning:

Thus tho’ causation be a *philosophical* relation, as implying contiguity, succession, and constant conjunction, yet 'tis only so far as it is a *natural* relation, and produces an union among our ideas, that we are able to reason upon it, or draw any inference from it. (T 1.3.6.16; SBN 94)

For Hume, the feature by virtue of which “we are able to reason upon” the causal relation is not the mere fact that the we “notice” casual relations in the absence of “special examination” (Penelhum), or that the “plain man” would “describe” causally related objects or events as “related” or “connected” (Bennett). Rather, it is the fact that there is a mental mechanism by which the mind is drawn from one idea to another that explains why we can “reason upon” the causal relation; in other words, causation’s status as a natural relation is constituted by the fact that it corresponds to this mental mechanism.

In other words, while Hume admittedly asserts in the quoted passage from T 1.1.5.1 that the qualities by which two ideas are connected together in the imagination (*viz*, resemblance, contiguity and causation) are just those qualities that common language takes to be “relations”, by his own lights he cannot but regard this fact as a feature of the natural/philosophical distinction that is explained by, and secondary to, the distinction between the association and the comparison of ideas that he describes in the immediately preceding passage.

A second standard interpretation of the natural/philosophical distinction agrees with the interpretation just considered that (as Bennett puts it) “any true statement of the form  $xRy$  ... asserts the holding of a philosophical relation between  $x$  and  $y$ ”, but goes on to define naturalness in terms of association rather than in terms of what plain men recognise as relations. Thus J. A. Robinson holds that “to say that a relation  $R$  is a ‘philosophical’ relation is to make a factually empty statement” because “*all* relations are philosophical”.<sup>5</sup> “*Naturalness*”, by contrast, “is then simply the property of any relation  $R$  between a thing or event  $A$  and a thing or event  $B$  (*not* between the *idea* of  $A$  and the *idea* of  $B$ ) whereby the observation of  $A$  and  $B$  standing to each other in the

relation *R* is enough to induce an association between the idea of *A* and the idea of *B*”; “naturalness ... is a purely contingent feature of a relation” (*ibid.*). Similarly, Harold Noonan says that the “term Hume uses for relations in general is ‘philosophical relations’”, while natural relations are “relations by which the human mind is naturally affected.”<sup>6</sup>

The difference between this interpretation and the one I want to defend is subtle but nonetheless important. According to the interpretation just considered, a natural relation is one that is also a principle of association – it induces a certain kind of mental activity in us – whereas a philosophical relation is simply a *relation*, whether or not it happens to induce that kind of mental activity. My preferred interpretation agrees with this interpretation when it comes to natural relations, but differs in that it conceives philosophical relations not as relations *simpliciter*, but as relations that correspond to a second kind of mental mechanism, namely the comparison of ideas.

What textual support can be garnered in favour of this latter interpretation, and against the former? According to the view I want to reject, we are to think of philosophical relations as simply the worldly relations: *any* relation between objects, or events, or whatever, will count as a philosophical relation. One reason to be skeptical about this view is that Hume is so confident that he has enumerated *all* the philosophical relations that he does not seriously consider whether there might be any philosophical relations that aren’t on his list. As he says:

It may perhaps be esteem’d an endless task to enumerate all those qualities, which make objects admit of comparison, and by which the ideas of *philosophical* relation are produc’d. But if we diligently consider them, we shall find that without

difficulty they may be compriz'd under seven general heads, which may be consider'd as the sources of all *philosophical* relation. (*T* 1.1.5.2; SBN 14)

According to the view under discussion – that the philosophical relations are simply *all* the relations there are – this is a spectacularly grand claim to make. We may “without difficulty” be able to figure out that all the philosophical relations *we know about* fall under the seven general headings; but why should there not be philosophical relations that fall outside the reach of possible human thought, or simply have not yet been discovered, which do not fall under one of these headings? What grounds can Hume possibly have for claiming that the “seven general heads” can be considered as the sources of *all* philosophical relations?

Of course, if we go for a really hard-line empiricist interpretation of Hume, the possibility of relations that lie outside the reach of possible human thought might not seem to be a problem, since we could attribute to Hume the view that it literally makes no sense to claim that, or even to wonder whether, there might be in-principle inaccessible relations. But this still leaves the possibility of philosophical relations that have not yet been discovered. Hume, on this view, can and should admit the possibility that there are such relations; why, then, doesn't he?

The interpretation I have offered, by contrast – according to which to designate a relation as “philosophical” is to make a claim about a certain kind of mental operation, *viz*, the comparison of ideas, as opposed to their association – explains this apparent omission. So long as Hume has conducted an exhaustive examination of the contents of the mind (which he clearly takes himself to have done), he *can* be confident that he has enumerated all the philosophical relations.<sup>7</sup> Any worldly relation that is

inaccessible or has yet to be discovered is not, *as things currently stand*, a philosophical relation: it is not a “subject of comparison” *for us*, and so plays no role in the operation of the mind. Thus when Hume says that the philosophical relations – which “make objects admit of comparison” – may be “compriz’d under seven general heads”, he really is solely interested in those qualities which *in fact* make objects admit of comparison.

This does, of course, raise the question why Hume divides relations into *two* kinds rather than three – for surely, on the proposed interpretation, he must in principle accept the possible existence of relations that are neither philosophical nor natural, *viz*, those that in fact, as things currently stand, play no role in human psychology.

(Analogy: suppose I either like or dislike anyone I meet on first encountering them.

Then I can divide *people I have met* into those whom I like and those whom I dislike.

But I do not thereby divide people *simpliciter* into just those two groups, since of course there are people whom I have not met, and whom I therefore neither like nor dislike.)

The answer, I believe, is that Hume is interested in the “science of man” (T

Intro 7; SBN xvi). He is attempting – and at this early stage in the *Treatise*, very

obviously so – to account for the operations of the human mind, and the possible existence of features of reality with which we have no acquaintance, and hence which

can play no role in our mental life, are simply irrelevant to his purposes. Thus, for

example, he is not even prepared to speculate about whether there are any objects

outwith the mind that correspond to our ideas, saying only that impressions of sensation

arise “in the soul originally, from unknown causes” (T 1.1.2.1; SBN 7). Our

impressions, and hence ideas, are what they are, whatever their provenance; Hume’s

interest is in how we construct our conception of reality, based on those resources. It is

therefore hardly surprising that when it comes to relations he should similarly be interested only in those relations – or rather, our representations thereof – that play a role in our thought.

It might be objected at this point that the defender of the view I am criticising – the view that the philosophical relations just *are* all the relations that exist, whether we know about them or not – can appeal to the limitations of Hume’s interest too: he does not consider philosophical relations that fail to play any role in our thought because they simply lie outwith his concerns. But that line of thought still fits badly with the text; Hume does after all say that the “seven general heads” may be “considered as the sources of *all*” philosophical relations (T 1.1.5.2; SBN 14, my italics), and not merely that they may be considered as the sources of all philosophical relations that we know about or of which we have ideas.

More generally, Hume *always* mentions philosophical relations in the context of the comparison of ideas; and, in particular, he says that resemblance “is a relation, without which no philosophical relation can exist; since *no objects will admit of comparison, but what have some degree of resemblance*” (T 1.1.5.3; SBN 14, my italics). So for Hume there cannot, in principle, be any philosophical relation that holds between two objects that do not in any sense resemble each other. Unless we take him to be making, on purely *a priori* grounds, the ontological claim that there is no worldly relation *x does not resemble y to any degree*, which would be a bad idea, his reason for holding that there can be no such philosophical relation must be the *a posteriori* psychological hypothesis that wholly non-resembling objects cannot be compared *by us*. Again we see an explicit connection between a relation’s being philosophical on the one hand and our ability to compare the objects so related on the other. Similarly, in the

two definitions of causation themselves, Hume says that we may consider causation “either as a *philosophical* or as a *natural* relation; either as a comparison of two ideas, or as an association betwixt them” (T 1.3.14.31; SBN 170). The correspondence between the philosophical/natural distinction and the comparison/association distinction could not be more clearly stated.

I conclude that Hume’s distinction between natural and philosophical relations is a purely psychological distinction: a relation’s status as natural or philosophical depends entirely on a distinction between two kinds of mental processes, namely the association and comparison of ideas. This does not, however, entail that relations *themselves* are purely psychological, any more than the claim that the distinction between people I like and people I dislike is a distinction grounded entirely in my psychology entails that people themselves are purely psychological entities. A relation that has yet to be discovered is, as things currently stand, neither natural nor philosophical, just as a person I have not yet met is, as things currently stand, neither a person whom I like nor a person whom I dislike.

### **3. The Procedural Interpretation**

Before applying the above account of natural and philosophical relations to Hume’s two definitions, I need to make an important but not especially controversial point about the distinctiveness of causation amongst the natural relations. Book I, Part 3 of the *Treatise* is driven by Hume’s interest in, first, reasoning from causes to effects and, second, the origin of the idea of causation. For Hume these are two sides of the same coin: it is only by understanding how we reason from causes to effects – the habit or custom that gets established through repeated observations of *As* being followed by *Bs*

– that we can understand the origin of the idea of cause, since the origin of a crucial component of the idea of cause, namely the idea of necessary connection, is a result of that very habit. One consequence of this is that in the case of causation, the psychological mechanism or transition of the mind that takes us from the idea of the cause to the idea of the effect (that is, causation considered as a natural relation) has a special function: it *generates causal judgements*. As Hume puts it in the first *Enquiry*:

But when one particular species of event has always, in all instances, been conjoined with another, we make no longer any scruple of foretelling one upon the appearance of the other, and of employing that reasoning, which can alone assure us of any matter of fact or existence. We then call the one object, *Cause*; the other, *Effect*. (EHU 7.27; SBN 75)<sup>8</sup>

How does all this bear on the two definitions? Well, one consequence of the account of natural and philosophical relations given above is that the distinction between causation considered as a natural relation and causation considered as a philosophical relation is a psychological distinction. This *can* amount to the difference described in §2: the difference between, on the one hand, “comparing” the ideas of, say, carbon consumption and global warming, and wondering whether or imagining that the two fall under the relation of causation (philosophical relation), and, on the other, imagining a toe-stubbing and the idea of pain immediately coming to mind (natural relation). But what Hume is interested in when it comes to the two definitions, I claim, is a difference in the kinds of mental operation by means of which we come to *formulate causal judgements*. In other words, the question he wants to answer is: how,

given that sensory input is restricted to past constant conjunction (and, sometimes, the impression of the cause) – there being no sensory impression of the relation *between* causes and effects – do we come by the judgement that one thing caused something else?

Here are the two definitions, as they appear in the *Treatise*:

There may two definitions be given of this relation, which are only different, by their presenting a different view of the same object, and making us consider it either as a *philosophical* or as a *natural* relation; either as a comparison of two ideas, or as an association betwixt them. We may define a CAUSE to be “An object precedent and contiguous to another, and where all the objects resembling the former are plac’d in like relations of precedency and contiguity to those objects, that resemble the latter.” If this definition be esteem’d defective, because drawn from objects foreign to the cause, we may substitute this other definition in its place, *viz.* “A CAUSE is an object precedent and contiguous to another, and so united with it, that the idea of the one determines the mind to form the idea of the other, and the impression of the one to form a more lively idea of the other”. (T 1.3.14.31; SBN 170)

Recall the distinction between philosophical and natural relations. A philosophical relation is “that particular circumstance, in which, even upon the arbitrary union of two ideas in the fancy, we may think proper to compare them”, while a natural relation is “that quality, by which two ideas are connected together in the imagination, and the one naturally introduces the other” (T 1.1.5.1; SBN 13). This, as I have said, is a

psychological distinction. Of course, since causation is both a philosophical and a natural relation, the “particular circumstance” by which we think proper to compare two ideas is *very same relation* as the “quality, by which two ideas are connected together”, namely, causation. But there are two distinct psychological procedures by which judgements or thoughts about causes and effects can be generated.

Let’s start with causation considered as a natural relation. I have observed a certain number of cases in which the contact of the white ball on the black ball has been followed by the black ball moving off. Next time I see the white ball hit the black, I automatically infer that the black will move, and, at the same time, come to think of the former as a cause of the latter (I “call the one object, *Cause*; the other, *Effect*” (EHU 7.27; SBN 75)). So I have come to make a causal judgement, and have done so as a result of the natural relation of causation: as a result, that is, of the operation of a particular associative mechanism.<sup>9</sup>

Now consider causation considered as a philosophical relation. Here’s an autobiographical example. Back when I was a graduate student, I sometimes found my hands shaking uncontrollably. Obviously my immediate assumption was that I had some terrible and incurable illness. But on reflection I rejected this hypothesis. I thought about past instances of uncontrollable hand-shaking, and considered whether I could draw any conclusions about what caused them. Could they have been caused by extreme nervousness or cold? No, because I could think of occasions when my hands had shaken uncontrollably but I was neither nervous nor cold. Could they, perhaps, have been caused by the spectacular quantities of coffee I would drink while passing the time in the coffee bar between seminars? Why, yes! I realised that on all the occasions I could recall, the hand-shakings had been preceded by high levels of coffee

intake. In other words, I considered these two ideas – high level of coffee intake and shaking hands – and “placed them” in the relation of causation, and I did so on the basis of contiguity, succession, and observed constant conjunction. (This, it turned out, was a very sensible thing for me to do. I cut down on the coffee and lo! The hand-shaking went away.) I made this judgement in the absence of a habit of inference from coffee-drinking to hand-shaking. Perhaps I hadn’t been paying enough attention to the contiguity of coffee-drinkings and hand-shakings. Maybe my inferential mechanisms aren’t as finely tuned as they could be. Either way, though, I compared two ideas, noted that the conditions of contiguity, succession and observed constant conjunction had been met, and formulated a causal judgement. That’s causation considered as a philosophical relation.<sup>10</sup>

The main piece of textual support for the Procedural Interpretation (in combination, of course, with the interpretation of the natural/philosophical distinction defended in §2) is the fact that Hume claims, clearly and explicitly, that the two definitions characterise causation separately as a philosophical and a natural relation. Indeed, as I shall argue in §4, it is the only interpretation that successfully deals with this piece of evidence. In addition, some contextual support accrues to the Procedural Interpretation from the role the two definitions play in Hume’s overall discussion of causal reasoning and the idea of necessary connection. Hume presents the two definitions as the culmination of this discussion. He says:

’Tis now time to collect all the different parts of this reasoning, and by joining them together form an exact definition of the relation of cause and effect, which makes the subject of the present enquiry. This order wou’d not have been excusable, of

first examining our inference from the relation before we had explain'd the relation itself, had it been possible to proceed in a different method. But as the nature of the relation depends so much on that of the inference, we have been oblig'd to advance in this seemingly preposterous manner, and make use of terms before we were able exactly to define them, or fix their meaning. We shall now correct this fault by giving a precise definition of cause and effect. (T 1.3.14.30; SBN 169)

The “reasoning” of which Hume is now collecting all the “different parts” is not merely the discussion of the origin of the idea of necessary connection that he has been conducting in T 1.3.14; he starts his enquiry into the idea of causation some twelve sections earlier in T 1.3.2, since when he has “beaten about” neighbouring fields including causal reasoning, unphilosophical probability, and so on. Seen in this light, Hume’s reference to the distinction between natural and philosophical relations, which immediately follows the paragraph quoted above, should seem entirely natural. For we have seen the distinction before in the discussion that he is now drawing together, in his discussion of causal reasoning.

Near the beginning of that discussion, he argues that it is “by EXPERIENCE only, that we can infer the existence of one object from that of another”, and he thereby “discovers” that constant conjunction is an additional relation (additional, that is, to contiguity and succession) between causes and effects (T 1.3.6.2; SBN 87). Here, Hume is very clearly talking about *observed* constant conjunction, as when he says: “Contiguity and succession are not sufficient to make us pronounce any two objects to be cause and effect, unless we perceive, that these two relations are preserv'd in several instances” (T 1.3.6.3; SBN 87); and “after the discovery of the constant conjunction of

any objects, we always draw an inference from one object to another” (*ibid.*). He then proceeds to present his famous discussion of causal reasoning or inductive inference, at the culmination of which he says:

Thus tho’ causation be a *philosophical* relation, as implying contiguity, succession, and constant conjunction, yet ’tis only so far as it is a *natural* relation, and produces an union among our ideas, that we are able to reason upon it, or draw any inference from it. (T 1.3.6.16; SBN 94)

I shall have more to say about this passage later on, but for now I want to make two points. First, Hume’s first *definition* of causation at T 1.3.14.31 is effectively repeating something he has said very much earlier on, at T 1.3.6.16: that causation as a philosophical relation implies contiguity, succession, and constant conjunction. This earlier mention comes at a point when talk of the *meaning* of “cause” is very much on the back burner; Hume’s concern in the preceding pages has been the nature of the *inference* from causes to effects (although of course it will turn out that the inference provides the impression-source for the idea of necessary connection). Second, the “constant conjunction” that Hume here refers to is most plausibly read as “observed constant conjunction”. Virtually all of Hume’s references to “constant conjunction” up to this point unambiguously refer to observed constant conjunction, and none unambiguously refer to universal constant conjunction. Hence Hume *cannot*, at T 1.3.6.16, be talking about the meaning of “cause”. Given the similarity between what Hume says at T 1.3.6.16 and what his two definitions say later on, it seems entirely

reasonable to read the two definitions in the same light; and this, of course, is just what the Procedural Interpretation recommends.

What about the fact that Hume says, clearly and explicitly, that the two definitions “fix the meaning” of cause? Isn’t that flatly incompatible with the Procedural Interpretation? Well, maybe so – but something has to give here. The two definitions cannot (as I argue in §4 below) simultaneously both “fix the meaning” of cause – at least in the contemporary philosophical sense of the words – *and* satisfy the other characterisation Hume immediately proceeds to give us: that the two definitions define causation first as a philosophical and second as a natural relation. Moreover, given that the two definitions are not equivalent, and given that there is no textual evidence that Hume intends “constant conjunction” in the first definition to mean “universal constant conjunction”, the claim that the two definitions “fix the meaning” of cause in anything like the standard contemporary sense of “meaning” is in big trouble in any case.

The Procedural Interpretation gets its main support, however, from the fact that other interpretations fail so badly in making sense of Hume’s claim to be characterising causation separately as a natural and a philosophical relation. I argue for this claim in §4 below. In the remainder of this section I consider some consequences of the Procedural Interpretation and a possible objection to it.

*(a) The meaning of “cause”*

The traditional interpretation of the two definitions is that they (or perhaps only the first) are genuine definitions in the contemporary sense; that is, that they capture the meaning of “cause”. The Procedural Interpretation denies this; indeed, it is completely

neutral with respect to the major interpretative positions on Hume on the nature of causation. When, for example, we form causal thoughts or judgements when comparing ideas – that is, as per the first definition – while we form those judgements on the basis of contiguity, succession and constant conjunction (that is, *observed* or *past* constant conjunction), what we judge to obtain between the ideas is *causation* – whatever that is – and not *merely* contiguity, succession and observed constant conjunction. The *idea* we deploy when we make the judgement is the idea *of causation*; and the first “definition” – characterising, as it does, merely the mechanism by which such a judgement is formed – is completely silent on what that idea involves. (I think that by his own lights Hume has already told us everything we need to know about what the idea involves; he hardly needs to repeat himself.)

Why, then, does Hume call the definitions “definitions”? Well, I think they *are* definitions of sorts, at least if we take a fairly broad view of what a “definition” could be. They are not definitions of “cause” *simpliciter*; rather they are, exactly as Hume says, definitions of “cause” *considered as* a philosophical and a natural relation. Thus considered, causation is a relation by means of which we get to perform certain sorts of mental operation. And the definitions specify what the two sorts of operation are: they are characterisations or specifications – and so “definitions” in at least a loose sense – of two kinds of mental operation, each of which delivers causal judgement.

Of course, this is not to say that it is merely a coincidence that the conditions mentioned in the definitions – contiguity, precedence, resemblance, constant conjunction and the determination of the mind – all play a role in Hume’s account of the composition of the idea of causation. In general, if you want to make a judgement about whether a given relation obtains, you’d better check, insofar as you are able, that

the conditions required for a pair of objects to fall under that relation really do obtain. In the case of causation considered as a philosophical relation, I need to check that the pair of objects meets the conditions of contiguity, succession, and observed constant conjunction, since these are all parts of the idea of causation. (This is not to say that the first “definition” is a definition in the traditional sense after all. Hume has already told us, repeatedly, that these are part of the idea of causation. What the first definition is *not* doing is telling us this one more time.) However, what I cannot do when considering two arbitrary ideas, as in the coffee case above, is consider whether they fall under the idea of necessary connection. Or rather, in considering whether to place them under the relation of causation, I am, precisely, considering whether to place them under the idea of necessary connection. And to do this, I consider whether they meet the conditions that necessarily-connected pairs of events have to meet.<sup>11</sup>

In the case of causation considered as a natural relation, by contrast, I don’t have to do any conscious considering at all; the causal judgement simply comes to me unbidden as a result of the relevant associative mechanism. In particular, I do not have to check whether the conditions that necessarily-connected pairs of events have to meet obtain, since the mere fact that the idea of the cause “determines the mind” to form the idea of the effect guarantees that the impression of necessary connection is present, and hence that I will judge the two events to be causally related.

*(b) The issue of extensional equivalence*

One virtue of the Procedural Interpretation is that the problem about the apparent lack of extensional equivalence between the two definitions drops off the radar all together. Hume is offering an account of two distinct mental operations by which causal

judgements are formed, and the fact that the mental operations are different in no way implies that there is any difference in the content of the idea of causation that those differently arrived-at judgements deploy.

We might, I suppose, ask: are the circumstances in which we deploy the method of the comparison of ideas precisely those in which the associative mechanism operates? But once the question is asked, it is obvious that the answer is “no”. As we have already seen, there can perfectly well be situations in which I come to form a causal judgement via the comparison of ideas, which are, precisely, not circumstances in which the associative mechanism is in a position to operate; the coffee case was just such a case. But this is not a problem for Hume. That the method of the comparison of ideas can be (and presumably normally, if not always, is) deployed in situations where the associative mechanism has not been set up is no cause for concern. Indeed, the very fact that our ability to form causal judgements is not restricted to those situations in which the relevant habit has been established underpins Hume’s “rules by which to judge of causes and effects” (see (d) below).

*(c) Why might the first definition be “esteem’d defective”?*

Hume tells us that the first definition may “be esteem’d defective, because drawn from objects foreign to the cause” (T 1.3.14.31; SBN 170). Different interpretations have different things to say about what he means by this; for example, Stroud says:

Hume thinks he has shown that it is only because things fulfil the conditions of the second “definition” that any things in the world are thought to be related *causally* or *necessarily* at all. We get the idea of necessary connection only because of the

passage of the mind from the thought of something to the thought of its “usual attendant”. That is perhaps why he feels constrained to include something like the second “definition” in any attempt to characterize our idea of causality.<sup>12</sup>

While I agree with everything Stroud says here apart from the last sentence, it is the last sentence that provides Stroud with an explanation for why the first definition might be thought to be defective in some way, and so I need to provide an alternative explanation.

Well, in fact things are a little more complicated, because, immediately after presenting the second definition, Hume says: “Shou’d this definition also be rejected for the same reason, I know no other remedy, than that the persons, who express this delicacy, should substitute a juster definition in it place” (T 1.3.14.31; SBN 170) – at which point he reminds us of some of the main points of his argument. So he appears to think that *both* definitions might be thought to be defective, and for the same reason.

We are heading into contentious territory here, since skeptical realist interpreters have taken Hume to be saying (as he does appear to say rather more clearly in the first *Enquiry*) that the two definitions really *are* defective, but that they cannot be remedied because penetration into the true nature of causation is impossible.<sup>13</sup> I shall return to this issue shortly; first, I shall explain why, according to the Procedural Interpretation, Hume might reasonably say that one or both of the definitions in the *Treatise* might *seem* (though in fact are not) defective.

A good proportion of Hume’s whole discussion of causal reasoning and necessary connection is aimed at refuting the claim that there is some detectable quality *in the cause* by virtue of which it guarantees its effect – and that, by virtue of detecting

it, we come to know that the effect will follow. Let's call this the "detectable-causes" view. Clearly Hume takes this to be a natural or popular view, since he goes to such great lengths to refute it. His central argument, of course, is that if there *were* such a quality, which we could observe, we would be able to infer the effect on first observing the cause – and manifestly we cannot do this. Now, what would the defender of the detectable-causes view want to say about the definitions of causation, considered as a philosophical and as a natural relation? Presumably they would want to say that, as far as the comparison of ideas is concerned, what we do when we consider two "arbitrary" ideas is that we inspect the first, discern the quality by which its effects are guaranteed, and then consider whether the second idea corresponds to one of those effects. Thus I might inspect my idea of excessive coffee-drinking and my idea of hand-shaking, discern that excessive coffee-drinking has the power to produce hand-shaking, and thus come to place the two ideas in the relation of causation. In the case of the association of ideas, they would want to say that the impression or idea of the cause reveals its power to produce the effect, and I thereby come to judge that the first is a cause of (and so will be followed by) the second.

These stories are wrong, Hume thinks, but insofar as the detectable-causes view is the view that he thinks many or most of his opponents are likely to have, he is sensitive to the kind of definition of causation *they* would want to see. And by their lights, Hume's definitions are both "defective, because drawn from something extraneous and foreign to the cause". Neither of Hume's definitions locates the source of causal judgement – whether via comparison or association – in any detectable power in the cause, although one might be tempted to think that the second definition is slightly less "defective" than the first, in that it at least identifies the observation of the

cause as a kind of trigger for the causal judgement: the thing that gets the associative mechanism to swing into action. The first definition does not even get us that much: there is nothing in the idea of the cause just by itself that provides a psychological basis for the causal judgement. The basis for causal judgement is merely the observed constant conjunction of pairs of contiguous events: that is, of features that are largely “extraneous and foreign” to the cause (that is, to the first object or event), since none of contiguity, succession and constant conjunction are features of the cause itself.

What if Hume thinks (as he seems to suggest in the *Enquiry*) that both definitions really *are* defective? Well, the Procedural Interpretation is, as I have said, neutral with respect to the broader interpretative positions on Hume on causation. If Hume really does think that “cause” picks out real but undetectable powers, as skeptical realist interpreters maintain, then the definitions are defective, in the sense that the psychological mechanisms that generate causal judgement are insensitive to the intrinsic feature of the cause that the judgement is a judgement about. If I want to figure out whether Josh has more apples than Charles – that is to say, if I want to decide whether to place the idea of the number of Josh’s apples and the idea of the number of Charles’s apples under the “greater than” relation – I do so by counting apples: the method by which I form the judgement involves detection of the very feature that the judgement is a judgement about. In the case of causation considered as a philosophical relation, by contrast, the method – ascertaining that contiguity, succession and past constant conjunction obtain – does not involve detection of the feature, namely a power of production, which the judgement is about. Similarly, when I make a causal judgement via the operation of the associative mechanism, I do so

because the associative mechanism itself generates the impression of necessary connection, and not because any necessitating power is detected in the cause itself.

So the Procedural Interpretation can make sense both of the claim that Hume's two definitions might *seem* defective but in fact are not, *and* of the claim, central to skeptical realist interpretations, that they really are defective. Not both at the same time, of course – rival interpretations of Hume's overall position on causation will take different views about which claim to ascribe to him.

*(d) Can the first definition do justice to what Hume says about causal reasoning?*

A possible worry emerges for my account of the two definitions when we consider what Hume says earlier on in the *Treatise*, right at the end of his discussion of causal reasoning:

Thus tho' causation be a *philosophical* relation, as implying contiguity, succession, and constant conjunction, yet 'tis only so far as it is a *natural* relation, and produces an union among our ideas, that we are able to reason upon it, or draw any inference from it. (T 1.3.6.16; SBN 94)

The reason why this – and indeed Hume's entire discussion of causal reasoning – might seem to be problematic for the Procedural Interpretation is that Hume *seems* to be saying that we can *never* come to believe some universal constant conjunction ("all *As* are *Bs*", say) unless we have observed for ourselves sufficiently many instances to have established the habit of inference from *As* to *Bs*. But if that is so, then causation considered as a philosophical relation, as I have described it, cannot be a route to causal

judgement – since precisely what happens in such cases is that I come to judge that *As* cause *Bs* (and hence that all *As* are *Bs*) in the absence of the operation of the associative mechanism.

I shall argue that Hume is clearly and explicitly committed to the claim that causation considered as a philosophical relation *is* a route to causal judgement. If that is a claim that he is not entitled to, in the light of the passage quoted above and the argument that precedes it, that is a problem for Hume himself, and not a problem for the Procedural Interpretation. However, I shall briefly sketch an interpretative line according to which he *is* entitled to hold that causation considered as a philosophical relation is a route to causal judgement, and according to which this claim is not after all in tension with his discussion of causal reasoning in T 1.3.6.

First, then, I need to show that Hume does hold that considering two arbitrary ideas under the relation of causation can indeed issue in causal judgement. Consider the fourth of his “rules by which to judge of causes and effects”: “when by any clear experiment we have discover’d the causes or effects of any phaenomenon, we immediately extend our observation to every phaenomenon of the same kind, without waiting for that constant repetition, from which the first idea of this relation is deriv’d” (T 1.3.15.6; SBN 173-4). Imagine (to modify the earlier example slightly) that I reflect on the circumstances under which I have suffered from shaking hands and discover two factors that have been present on each occasion: coffee and carrot cake. So I cannot infer that coffee is a cause of my hand-shaking. I therefore devise a “clear experiment” – I drink a lot of coffee but desist from the cake – and my hands begin to shake. I conclude that coffee-drinking is indeed a cause of the hand-shaking – not just in this case, but in past cases too; and indeed in any future cases where I foolishly drink too

much coffee. That is, I extend my observation to every phenomenon of the same kind. And, by hypothesis, I do so without “waiting” for the constant repetition that would establish the inferential mechanism.

Of course, the principle that I am deploying here – “[t]he same cause always produces the same effect, and the same effect never arises but from the same cause” (T 1.3.15.6; SBN 173) – is one that, as Hume says, “we derive from experience” (*ibid.*). But this is not to say that we can somehow shoehorn one-shot experiments into the model of associative reasoning. Rather, it is merely to reiterate that the principle is based on our experiences of constant conjunction: it is not knowable *a priori*.

Several of the other rules also commit him to thinking that causal judgement can be had in the absence of association. For example, here is the sixth rule:

The difference in the effects of two resembling objects must proceed from that particular, in which they differ. For as like causes always produce like effects, when in any instance we find our expectation to be disappointed, we must conclude that this irregularity proceeds from some difference in the causes. (T 1.3.15.8; SBN 174)

In other words, if we have observed that all *As* have been followed by *Bs* in our experience, and then come across an *A* that is not followed by a *B*, we must conclude that there is some hidden feature of the previous *As* – call it *F* – that was the underlying cause of the *Bs*, such that the last *A* we saw, the one that confounded our expectations, failed to have *F*. So in fact there are two kinds of object or event, *A1* and *A2*, such that *A1s* cause *Bs* but *A2s* don’t (and we can’t observationally tell the difference between

*A1s* and *A2s*). Thus we reach a causal conclusion – that *A1s* cause *Bs*, while *A2s* don't – but not solely on the basis of the associative mechanism. For as far as the associative mechanism is concerned, the *A1s* and the *A2s* are alike (they are all *As*); and since we no longer have observed constant conjunction between *As* and *Bs*, the associative mechanism will presumably shut off so that future exposure to *As* will, just by itself, deliver no causal judgement. Hume's rules, then, provide a basic scientific method; and it is a scientific method that delivers causal judgement in the absence of the operation of the associative mechanism in the particular case under consideration.

Hume later reiterates his point that differences in effects must proceed from differences in their causes in the context of arguing that “our actions have a constant union with our motives, tempers, and circumstances” (T 2.3.1.4; SBN 401) and therefore are subject to causal necessity just as much as is the behaviour of material objects. In cases where observed constant conjunction is lacking:

The mind ballances the contrary experiments, and deducting the inferior from the superior, proceeds with that degree of assurance or evidence, which remains. Even when these contrary experiments are entirely equal, we remove not the notion of causes and necessity; but supposing that the usual contrariety proceeds from the operation of contrary and conceal'd causes, we conclude, that the chance or indifference lies only in our judgment on account of our imperfect knowledge, not in the things themselves, which are in every case equally necessary, tho' to appearance not equally constant or certain. (T 2.3.1.12; SBN 403-4)

When the “contrary experiments are entirely equal”, then – that is, if in our experience *As* have been followed by *Bs* exactly half the time – still we come (or at least *can* come) to formulate causal judgements; again, in this case the judgement will (or should) be that there are “concealed causes” in half of the *As*, whose presence causes *Bs*. If we were to rely on the associative mechanism alone, no such judgement would be possible, since by hypothesis there is no observed constant conjunction that would be capable of delivering it.

The point here is that in conceiving causation considered as a philosophical relation as a route to causal judgement, the Procedural Interpretation merely attributes to Hume a claim to which he commits himself to in any case: the claim that causal judgements can be made in the absence of the operation of the associative mechanism. Nonetheless, this claim does itself appear to be in tension with Hume’s claim that “tho’ causation be a *philosophical* relation, as implying contiguity, succession, and constant conjunction, yet ’tis only so far as it is a *natural* relation, and produces an union among our ideas, that we are able to reason upon it, or draw any inference from it”. One might take this tension to constitute a reason to reject the Procedural Interpretation, since one might take this passage (and the argument that precedes it) to be so central to Hume’s overall position that his later apparent claims to the effect that we *can* “reason upon” causation considered as a philosophical relation should be taken as something to explain away as a mere infelicity on Hume’s part.

This line of thought strikes me as implausible, given the centrality of Hume’s argument that the unpredictability of human actions is no grounds for denying that they are subject to causal necessity to his argument for the compatibility of liberty and necessity (which is repeated in the first *Enquiry*). It would therefore be better if we

could reconcile Hume's apparent insistence that we cannot "reason upon" causation considered as a philosophical relation with the reasoning that he later both endorses in his "rules by which to judge of causes and effects" and deploys in his discussion of free will.

I believe that we *can* do better. The basic idea is just this: once we have the idea of causation – an idea that requires that we have the habit of association that generates the needed impression-source for the idea of necessary connection – we are perfectly capable of deploying it in cases where the associative mechanism does not operate (that is, "without waiting for that constant repetition, from which the first idea of this relation is deriv'd"), just as we can for any other idea. Once I have had an impression of red, and thus have the idea of red at my disposal, I can form the belief that Peter has bought a pair of red gloves on the basis of testimony, without having to see them for myself; I don't have to have the impression of red *right now* in order to be able to deploy the idea. Similarly, having considered the matter and noted that previous hand-shakings have invariably followed shortly after excessive coffee-drinking, I can deploy the idea of necessary connection and come to judge that the latter were causes of the former – and thereby judge that (unless I cut back on the coffee) they will continue to be in the future. And so I come to believe, perfectly legitimately, that future excessive coffee-drinkings and hand-shakings will also be constantly conjoined.

So we need the inferential mechanism if we are to form any beliefs about the unobserved whatsoever; but, thanks to the convenient fact that the inferential mechanism gives rise to the impression of necessary connection, which is what allows me to have the concept of causation, I can then deploy that concept in the absence of the operation of the mechanism. So it is true that it is "only so far as [causation] is a

*natural* relation, and produces an union among our ideas, that we are able to reason upon it, or draw any inference from it”, in the sense that the associative mechanism is required for causal reasoning *in general*;<sup>14</sup> nonetheless, cognitively sophisticated creatures that we are, we can transcend the limited scope of the associative mechanism and additionally come to form causal judgements on the basis of the comparison of ideas.

Putting this thought in the context of Hume’s discussion of causal reasoning, he says very near the beginning:

Thus in advancing we have insensibly discover’d a new relation betwixt cause and effect, when we least expected it ... This relation is their CONSTANT CONJUNCTION. Contiguity and succession are not sufficient to make us pronounce any two objects to be cause and effect, unless we perceive, that these two relations are preserv’d in several instances. We may now see the advantage of quitting the direct survey of this relation, in order to discover the nature of that *necessary connexion*, which makes so essential a part of it. There are hopes, that by this means we may at last arrive at our propos’d end; tho’ to tell the truth, this new-discover’d relation of a constant conjunction seems to advance us but very little in our way. For it implies no more than this, that like objects have always been plac’d in like relations of contiguity and succession; and it seems evident, at least at first sight, that by this means we can never discover any new idea ... (T 1.3.6.3; SBN 87-8)

Hume’s discussion of causal reasoning then proceeds without any mention of necessary connexion, since he needs to identify the associative mechanism before he can proceed

to identifying the impression-source of the idea of necessary connection. But note that necessary connection does something that constant conjunction cannot do. Constant conjunction (by which Hume means *observed* constant conjunction) does indeed imply no more than that like objects *have* always been placed in like relations of contiguity and succession. It therefore cannot, just by itself, bridge the inferential gap between cause and effect, since it fails to imply anything about what we have not yet observed. Necessary connection, by contrast, *does* bridge that gap: to judge that *As* cause *Bs* just *is, inter alia*, to judge that future *As* will be followed by *Bs*. Thus, I claim, our ability to deploy the idea of necessary connection – an ability we will possess if we have the associative mechanism that generates its impression-source – allows us to bridge the inferential gap in cases where the associative mechanism is not operating.

This is of course extremely quick, but my aim here is only to sketch a way in which the two aspects of Hume's thought that are apparently in tension might be reconciled. The basic point remains even if the sketch fails to convince: the claim that the comparison of ideas can issue in causal judgement is central to the Procedural Interpretation, but there is plenty of textual evidence that it is a claim to which Hume does indeed subscribe. That is a problem for Hume if the tension cannot be resolved; but it is not a problem for the Procedural Interpretation in particular.

*(e) Psychology and epistemology*

As is often noted, Hume is not very careful to distinguish psychological issues from normative issues. In keeping with the more overtly psychological tone of the *Treatise*, I have been presenting the Procedural Interpretation in psychological terms: in terms of the mental processes by which, in fact, human beings come to form causal judgements.

But, as always, normative questions have not been far away – particularly in the discussion above about Hume’s rules. The rules seem to tell us how we *ought* to proceed in our theorising about causation (hence “rules”), and not simply how we do, in fact, proceed.

I have proceeded, then, as though the two definitions are broadly empirical claims about the mechanisms by which causal judgements are, in fact, formed in human beings; but that cannot be exactly right. For one thing, there are plenty of other ways of coming to make causal judgements – indoctrination, for example – and Hume thinks many of them are *bad* methods: the maxims learned through “education”, he says, “are frequently contrary to reason, and even to themselves in different times and places” (T 1.3.9.19; SBN 117). For another thing, both of the central methods – association and comparison – can go wrong. The associative mechanism that delivers causal judgements is one that could be defective in various different ways in different people: it could kick in after only a couple of repetitions, say, or it could be absent all together (though doubtless someone who lacked the mechanism wouldn’t last long). And someone who has casually observed a single instance of successive events (turning on the kettle and the neighbour’s light coming on, say) might come to judge on the basis of the comparison of ideas that they are causally related. Hume does not and should not say that, as a matter of psychological fact, we never go wrong.

Fortunately it is a fairly easy matter for the Procedural Interpretation to accommodate a rather more normative angle on Hume’s concerns, for we can think of the definitions as specifying how properly functioning, epistemically responsible people get to form their causal judgements. Someone who judged on the basis of the comparison of ideas that their turning on the kettle caused their neighbour’s light to

come on would be failing to conduct their mental life appropriately: their failing would be a normative one. Someone who had a faulty associative mechanism would similarly be failing to meet appropriate normative standards – though of course this would be due to a psychological defect that they presumably could not do anything about. We can think of the definitions as simply implicitly excluding such people from their range of application.

This kind of approach is, I think, one that we are forced to take at various other points in Hume's writing in any case. For example, in his account of knowledge, Hume applies a conceivability test: it is the fact that we can conceive that the impact of the first billiard ball fails to be followed by movement of the second that tells us that the expectation of the movement cannot be an item of knowledge, whereas our inability to conceive of two being greater than or equal to four tells us that  $2 < 4$  *is* an item of knowledge. Well, some people presumably have vastly impoverished imaginations and fail to be able to conceive all kinds of things that the rest of us find easy to imagine. Is the expectation that the second ball will move an item of knowledge *for them*, according to Hume? Or are such people counter-examples to the claim that what can be known is that whose opposite cannot be conceived? No, surely not. Hume is only talking about properly-functioning human beings whose imaginative powers are in full working order. Similarly, someone who makes a causal judgement via the comparison of ideas but who neglects to check that the contiguity requirement is met, or who can clearly recollect a counter-example to the constant conjunction requirement but ignores it, is not a counter-example to Hume's definition of causation considered as a philosophical relation, because that definition, while a claim about human psychology, incorporates an implicit normative element. Basing one's causal judgements on

contiguity, succession, and observed constant conjunction is what *properly-functioning, epistemically responsible* people actually do. So it is a claim about human psychology, but a claim that applies only to those people whose psychology is in good working order.<sup>15</sup>

This is not to say, however, that we can now see the two definitions as providing accounts of the circumstances under which causal judgements are *true*, thus bringing the Procedural Interpretation much closer to the conception of the definitions as genuine “definitions”; the suggestion here is not that we restrict our “properly-functioning” subjects to what Don Garrett calls “idealised spectators”.<sup>16</sup> Properly-functioning, epistemically responsible people can get things wrong: they can, through no fault of their own, observe unrepresentative samples, forget counter-examples to the observed constant conjunction requirement, and so on. Since the definitions are not, on the Procedural Interpretation, supposed to be definitions of *causation* (as opposed to specifications of the mechanisms by which causal judgements are formed), the fact that those mechanisms can deliver judgements that are in fact false even in properly-functioning human beings presents no particular problem for the interpretation, or indeed for Hume himself.

#### **4. Existing interpretations and their failings**

There have been almost as many interpretations of what Hume is doing in his two definitions as there have been commentators on them, and of course an exhaustive survey is impossible. However, so far as I can tell, virtually all existing accounts of the two definitions fall into one of the four groups identified below<sup>17</sup> (and most of them fall into the first group). I shall argue that none of the three kinds of view makes sense of

Hume's claim to be defining causation first as a philosophical relation and second as a natural relation.

*(a) The first definition as a genuine definition*

Perhaps the most common interpretation of the two definitions is that the first definition tells us what causation *is* – it really is a definition – whereas the second says something non-definitional about what goes on in our minds. For example, Harold Noonan says: “The view of the object [as in ‘two views of the same object’] provided by the definition of *cause* as a philosophical relation is a view of it as it is in itself ... The view of the object presented by the definition of *cause* as a natural relation is a view of it in its role as something which affects the mind in a certain way”.<sup>18</sup> And George Dicker says that the first definition “is supposed to be the complete definition of causation, as it exists objectively in nature. [The second definition], by contrast, makes only a contingent claim about *observed* cases of causation”.<sup>19</sup>

Does this interpretation of the two definitions make sense of Hume's claim to be defining causation first as a philosophical relation and, second, as a natural relation? One way to make the case that it does would be to appeal to the view of philosophical relations according to which they are, simply, worldly relations,<sup>20</sup> so that a definition of cause as a philosophical relation really would say what conditions two objects or events have to satisfy in order for it to be the case that the first caused the second. As Robinson puts it, “the cause-effect relation, being a relation, is *ipso facto* a philosophical relation, and therefore to define it ‘as’ a philosophical relation is, simply, to define it”.<sup>21</sup>

I have already argued in §2, however, that this is not a plausible view to take of Hume's conception of a philosophical relation. If we think of Hume's conception of a philosophical relation as a "particular subject of comparison", that is, as a relation that plays a certain role in the operation of the mind, then the first definition – of causation "considered as" a philosophical relation – cannot simply be a definition of *cause* "as it is in itself", as Noonan puts it. For to consider a relation *as* a philosophical relation is *not* to consider it "as it is in itself", but rather to consider the particular role it plays in our mental lives, namely its role in the comparison of ideas.

One might still think that the kind of view described above, and in particular the view that the first definition describes causation as it is "in the objects", can be made to fit the fact that Hume's interest in causation as a philosophical relation is an interest in the comparison of ideas. After all, one might think that the conditions that are required in order for us to judge, via the comparison of ideas, one event to cause another just *are* those "features of events contributed by the natural world", *viz*, contiguity, succession, and constant conjunction – in which case, the first definition could simultaneously be *both* an account of how we judge (via the comparison of ideas) two events to be causally related *and* an account of what causation "in the objects" really is.

Unfortunately this cannot be right. Any conception of the first definition as an account of causation "in the objects" *must* read "constant conjunction" as "*universal* constant conjunction", since of course events that have merely been conjoined in my experience might easily fail to be universally constantly conjoined and hence fail to be causally related. But causal judgement that is delivered by the comparison of ideas operates on the basis of *observed* constant conjunction. In other words, Hume holds that coming to judge two events to be causally related – on the basis of observed

constant conjunction – is psychologically *prior* to judging them to be universally constantly conjoined.

To see why this is so, consider Hume’s discussion of causal reasoning. How is it, Hume asks, that we can come to believe a universal constant conjunction? It is via the relation of cause and effect: the “only connexion or relation of objects, which can lead us beyond the immediate impressions of our memory and senses, is that of cause and effect” (T 1.3.6.7; SBN 89). So we do not *first* come to believe some universal constant conjunction, and *then* infer that events of the first kind are causally related to events of the second kind. Rather, we judge, on the basis of *observed* constant conjunction, that the two kinds of event are causally related, and *thence* infer that the two kinds of event are universally constantly conjoined. (Recall the coffee case: I placed the ideas of coffee-drinking and hand-shaking in the relation of causation on the basis of *past* constant conjunction, and *thence* inferred that, if I carried on drinking too much coffee, the hand-shaking would persist.) So the first definition, if it is to be an account of causation “in the objects”, cannot simultaneously be an account of how we judge two events to be causally related, since our judgements are based on observed constant conjunction, whereas an account of causation as it is “in the objects” must take the constant conjunction to be universal.

The view under consideration in this section cannot, therefore, be correct. If the account of philosophical relations of §2 is correct, then to consider causation as a philosophical relation is not to consider causation as it is in itself, as opposed to considering it in the context of the comparison of ideas. Nor can the first definition simultaneously describe causation as it is in itself *and* as the basis for causal judgement arrived at via the comparison of ideas, since the basis for causal judgement is *observed*

constant conjunction, while causation as it is in itself is, or implies, universal constant conjunction.

*(b) Circumstances of belief acquisition*

A second way of conceiving the definitions, which does not take the first definition to be an account of causation as it is “in the objects”, is put forward by Edward Craig, who says that the definitions characterise the “circumstances under which belief in a causal connection arises, one concentrating on the outward situation, the other on the state of the believer’s mind that those outward facts induce”.<sup>22</sup>

This line of interpretation is considerably closer to the Procedural Interpretation: it denies that Hume is *defining* causation in the standard contemporary sense of “definition”, and it asserts that Hume’s interest in the two definitions is in the formation of causal judgements. It differs from the Procedural Interpretation, however, in effectively ignoring causation considered as a philosophical relation all together. We are supposed to read the first definition as “concentrating on the outward situation” that induces the transition of the mind, while the second definition concentrates on the “observer’s state of mind”. On this interpretation, then, *both* definitions are focussed squarely on causation considered as a natural relation – but from the perspectives of the outward situation and the observer’s state of mind respectively. And of course the interpretation thereby fails to do justice to Hume’s claim that the first definition characterises causation considered as a *philosophical* relation.

I argued in §3(d) above that Hume commits himself to the claim that association is not the only mechanism by which causal judgements can arise. If that is right, then it provides an additional argument against Craig’s interpretation. If the two

definitions are, as Craig rightly contends, supposed to capture the “circumstances under which belief in a causal connection arises”, then to characterise *both* definitions as characterising the circumstances in which the associative mechanism operates not only falls foul of the natural/philosophical distinction; it also implicitly holds Hume to the view that belief in causal connection (or, as I would prefer to put it, the formation of a causal judgement) can *only* arise in those circumstances. I believe there is plenty of textual evidence that Hume holds no such view.

*(c) Defining causation in a loose sense*

According to Simon Blackburn, the “first ‘philosophical definition’ describes the contribution of the world, insofar as we can apprehend it, and the second ‘natural’ definition describes the nonrepresentative, functional difference in the mind that apprehends the regularity”.<sup>23</sup> Similarly, according to Angela Coventry the “first definition points to features of events contributed by the natural world, and the second definition records a spectator’s reaction or response to those events”<sup>24</sup>; and according to Garrett “we can define ‘cause and effect’ either in terms of the *constant conjunction* that in fact produces the determination or transition ... or we can define ‘cause and effect’ in terms of the *association and inference*’ (1997, 106).<sup>25</sup> I take the common claim here to be that the two definitions ‘define’ causation in a fairly loose sense, in that they tell us something about how the *idea* of causation comes about.<sup>26</sup> While similar to Craig’s interpretation in that it shares the thought that the first definition specifies the contribution of the world and the second specifies the contribution of the mind, on the current proposal the different contributions are contributions to the

formation of the *idea* of causation, rather than (as Craig has it) the circumstances under which particular causal *judgements* are formed.

Does the “definition in a loose sense” conception of the two definitions do justice to the natural/philosophical relation distinction? No, it does not. The distinction between natural and philosophical relations is not a distinction between the contribution of the world and the contribution of the mind to the acquisition of the idea of a given relation, since it is not a distinction that concerns the *acquisition* of ideas of relations at all. Rather, it concerns the *deployment* of ideas. In particular, causation cannot function as a philosophical relation for someone who does not already possess the idea of causation. One cannot acquire the idea of causation via the comparison of ideas; that is one of Hume’s points in his earlier discussion of causal reasoning. Hence no interpretation that characterises the first definition as an explanation of the circumstances under which the idea of causation is *acquired* can succeed in doing justice to Hume’s claim that the first definition characterises causation considered as a philosophical relation.

*(d) The inadequate-idea interpretation*

A third interpretative move has been to think of the two definitions not as defining causation itself, but of describing our *idea* of causation – the thought being that our idea of causation is inadequate to its object, namely real causal power. (On this view, this inadequacy of the idea is why Hume thinks his definitions may be “esteem’d” – and indeed are – “defective”.) Thus John Wright says that the first definition “is based on our actual sense-experience of objects and shows the limits of the objective ‘meaning’ we can assign to causality ... Hume’s definition is presented as one which

must serve in lieu of insight into real objective causal connection”.<sup>27</sup> And Galen

Strawson says:

Roughly speaking, the two definitions of cause give an account of the content of the *idea*'s impression-sources. They give information about what we positively-contentfully mean, and indeed about all we can really (positively-contentfully) mean, according to the theory of ideas, when we talk about causes. But they are seriously imperfect as definitions (and whatever exactly Hume thought the purpose of definitions was), because drawn from something “foreign to the cause” itself. And so they cannot really capture the true nature of causation at all.<sup>28</sup>

My question, again, is whether such a conception of the two definitions can make sense of the natural/philosophical distinction; and my answer – again – is “no”.<sup>29</sup> If both definitions are descriptions of the content of the idea of causation, it is unclear why Hume would even feel the need to provide *two* definitions in the first place. Couldn't he just list contiguity, succession, constant conjunction and necessary connection in a single definition, and leave it at that? Or, to put the point another way, Hume surely thinks that we have *one* idea of causation (whether defective or not), which we deploy every time we think or make judgements about causation. The difference between natural and philosophical relations cannot explain why, when considered as a philosophical relation, the content of the idea needs to be described one way, and when considered as a natural relation it needs to be described differently. *What we think* when we make a causal judgement remains constant, whatever the mechanism or method by which we come to make it.

The existing accounts of the two definitions thus fail to do justice to Hume's own claim that they characterise causation as a philosophical and a natural relation – once we have a proper understanding of what the natural/philosophical distinction amounts to. I therefore recommend the Procedural Interpretation on the grounds that it *does* make sense of Hume's claim.

### **5. The *Treatise* and the *Enquiry***

My discussion of the two definitions has been exclusively concerned with the *Treatise* versions of them. However, while the definitions appear in roughly the same form in the first *Enquiry*, there are some salient differences between the *Treatise* and the *Enquiry* – both in the formulation of the definitions and elsewhere – that require attention.

In the *Enquiry*, Hume presents his two definitions as follows:

Similar objects are always conjoined with similar. Of this we have experience. Suitably to this experience, therefore, we may define a cause to be *an object, followed by another, and where all the objects similar to the first are followed by objects similar to the second*. Or in other words, *where, if the first object had not been, the second never had existed*. The appearance of a cause always conveys the mind, by a customary transition, to the idea of the effect. Of this we also have experience. We may, therefore, suitably to this experience, form another definition of cause, and call it, *an object followed by another, and whose appearance always conveys the thought to that other*. (EHU 7.29; SBN 76-7)

It's worth saying at the outset something about how this reads according to the Procedural Interpretation. The basic idea is simply to read "suitably to this experience, we may define a cause to be ..." as something like, "the circumstances, as represented by our ideas, under which we come to form a causal judgement are ...". In the first case, we consciously ascertain by reflection that the italicised conditions are met, and thereby, via the comparison of ideas, come to judge that the first object caused the second. In the second case, it is the associative mechanism that generates the causal judgement.

What are the relevant differences between the *Treatise* and the *Enquiry*? First, in the *Enquiry* Hume provides an alternative version of the first definition: "where, if the first object had not been, the second never had existed". It also drops the contiguity requirement. Since neither of these changes is relevant to the viability of the Procedural Interpretation as it applies to the *Enquiry*, I shall not comment on them here.<sup>30</sup>

Second, Hume appears to be considerably less equivocal about the adequacy of the definitions in the *Enquiry*. I have already discussed this issue in §3(c) above.

Third, my claim in §4 that the first definition in the *Treatise* should be read as referring to past or observed constant conjunction gains some support from the considerably less ambiguous wording Hume uses in the *Enquiry* version. For what we have "experience" of is not universal constant conjunction – we very rarely have experience of *that* – but observed (that is, experienced) constant conjunction only. Independently of the viability of the Procedural Interpretation, this should give those commentators who give "constant conjunction" in the first definition a universal reading rather more pause for thought than it appears to have done.

Fourth, and most importantly in the context of the argument of this paper, explicit mention of the distinction between philosophical and natural relations is entirely absent from the *Enquiry*. And this of course raises the possibility that, whatever the merits of the Procedural Interpretation as applied to the *Treatise*, it cannot be applied to the first *Enquiry*.

On the other hand, Hume's account of the operation of the human mind, as put forward in the early pages of the *Enquiry*, while hugely curtailed, does not appear to disagree significantly with what he says at the beginning of the *Treatise*. In particular, the natural relations – contiguity, resemblance, and cause and effect – appear under the name of “principles of connexion” or “principles of association” (EHU 3.2; SBN 24). And Hume's (very brief) account of *a priori* knowledge still divides the “objects of human reason” into “relations of ideas” and “matters of fact”, so at least some of Hume's philosophical relations (those that generate *a priori* knowledge via the comparison of ideas) get a passing mention. So while the expressions “philosophical relation” and “natural relation” do not appear, the basic distinction between them would still seem to be implicit – in which case, the lack of appearance of the particular terms does not seem to be sufficient to render the Procedural Interpretation inapplicable.

Moreover, given that the *Treatise* version of the two definitions is supposed to characterise causation first as a philosophical relation and second as a natural relation, it is unclear why Hume would persist in offering two definitions of cause in the *Enquiry* if he no longer thought of himself as characterising causation in that way. If these two definitions are not supposed to be accounting for causation separately as a philosophical and as a natural relation – however we cash that out – why are there *two*? And why, if the definitions are supposed to be doing something different in the *Enquiry*

to what they do in the *Treatise*, are they so similar? In the absence of an adequate answer to these questions, it seems that we should think of the *Enquiry* definitions as playing the same role as they do in the *Treatise*, that is, as characterisations of causation as a philosophical and a natural relation.

Despite some relevant differences between the *Enquiry* and the *Treatise* formulations of the two definitions, then, there is no great barrier to transposing the Procedural Interpretation of the two definitions as they appear in the *Treatise* onto the *Enquiry* version – although of course the major item of textual support (the claim that the two definitions characterise causation considered separately as a philosophical and a natural relation) is only to be found in the *Treatise*.

There are, of course, those Hume scholars – in particular Galen Strawson – who insist that the *Enquiries* are to be taken as Hume’s considered and best-expressed view, and that the *Treatise* should be read in that light.<sup>31</sup> This is a controversial position, but someone who holds it might agree with Penelhum’s judgement that the presence of the natural/philosophical distinction in the context of the two definitions in the *Treatise* is an “obfuscatory complication”, the absence of which, in the first *Enquiry*, is a positive “benefit”. The argument of this paper, however, has been that, properly understood, neither the natural/philosophical distinction itself nor the role it plays in the two definitions is at all “obfuscatory”. This places a dialectical burden on any defender of the Strawsonian line on the way in which the *Treatise* should be read who wishes to deny that the two definitions should be interpreted in the light of the natural/philosophical distinction. Such an interpreter needs to explain either why Hume would later wish to reject the distinction, or perhaps exactly what Hume’s intended, but badly expressed, intention was in drawing the distinction in the first place. And they

also need to explain why Hume should feel the need to provide *two* definitions of causation, given that the most obvious explanation – that he intends to “define” causation considered separately as a philosophical and a natural relation – is not available.

## NOTES

---

<sup>1</sup> Terence Penelhum, *Hume* (London: Macmillan, 1974), 54.

<sup>2</sup> References to the *Treatise* are to David Hume, *A Treatise of Human Nature*, ed. David Fate Norton and Mary J. Norton (New York: Oxford University Press, 2000), hereafter cited as “T” followed by Book, part, section, and paragraph numbers; and to *A Treatise of Human Nature*, ed. L. A. Selby-Bigge, revised by P. H. Nidditch, 2nd ed. (Oxford: Clarendon Press, 1978), hereafter cited as “SBN” followed by page number.

<sup>3</sup> Jonathan Bennett, *Locke, Berkeley, Hume* (Oxford: Clarendon Press, 1971), 250-1.

<sup>4</sup> Penelhum, *Hume*, 54.

<sup>5</sup> J. A. Robinson, “Hume’s two definitions of ‘cause’ ”, *Philosophical Quarterly* 12 (1962): 162-71, 166.

<sup>6</sup> Harold Noonan, *Hume on Knowledge* (London: Routledge, 1999), 73.

<sup>7</sup> Well, perhaps he shouldn’t be *quite* as confident as he is. After all, his enumeration appears to have been done on the basis of a bit of casual armchair psychology.

<sup>8</sup> References to the first *Enquiry* are to David Hume, *An Enquiry Concerning Human Understanding*, ed. T. L. Beauchamp (New York: Oxford University Press, 2000), hereafter cited as “EHU” followed by section and paragraph numbers; and to Hume, *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, ed. L. A. Selby-Bigge, revised by P. H. Nidditch, 3rd ed. (Oxford: Clarendon Press, 1975), hereafter cited as “SBN” followed by page number.

<sup>9</sup> Note that causation “considered as” a natural relation need not *always* issue in causal judgement. When I am thinking about my friend and my mind “naturally” passes to the idea of her mother – a cause of her existence – I do not thereby come to *judge* that her

---

mother is a cause of her existence, since of course I already formed that judgement earlier on (which is precisely why my mind makes the transition it does). The claim here is merely that causation considered as a natural relation *can* (and frequently does) issue in causal judgement, and not that it *always* does.

<sup>10</sup> A standard view about Hume's story here, which I am implicitly denying, is that the route to causal judgement proceeds from observed regularity (all observed *F*s have been followed by *G*s) to universal generalisation (all *F*s are *G*s) and thence to the causal judgements (*F*s cause *G*s). In other words, according to this view enumerative induction precedes causal judgement. But this story ignores Hume's central contention that the "only connexion or relation of objects, which can lead us beyond the immediate impressions of our memory and senses, is that of cause and effect" (T 1.3.6.7; SBN 89). I say a little more about this in §4(a) below and in my *Hume on Causation* (Abingdon: Routledge, 2007), §3.8.

<sup>11</sup> The reader may of course plug in their preferred account of what Hume thinks it is for two events to be necessarily connected here.

<sup>12</sup> Barry Stroud, *Hume* (London: Routledge, 1977), 90.

<sup>13</sup> See Galen Strawson, *The Secret Connexion* (Oxford: Oxford University Press, 1989), 208; Stephen Buckle, *Hume's Enlightenment Tract* (Oxford: Oxford University Press, 2001), 207-8; and John Wright, "Hume's Causal Realism", in R. Read and K. Richman (eds), *The New Hume Debate* (London: Routledge, 2000), 91.

<sup>14</sup> I say more about this in my *Hume on Causation*, Chapter 3.

<sup>15</sup> Of course, such people can also form causal judgements on the basis of testimony, so long as their source is reliable. Hume's definitions are only intended to capture the

---

mechanisms that deliver causal judgements *ex nihilo*: causal judgements that are, as it were, made by oneself and not merely transmitted from some other source.

<sup>16</sup> Don Garrett's interpretation of the two definitions (*Cognition and Commitment in Hume's Philosophy* (Oxford: Oxford University Press, 1997, Chapter 4) bears some similarities to the Procedural Interpretation (though see §4(b) for what I think is a fatal objection to his proposal). However, he uses the notion of an "idealised spectator" to retrieve from the two definitions something reasonably close to "definitions" in the contemporary sense. Having jettisoned the idea that the definitions are "definitions" in anything like this sense, I see no reason to follow suit and think in terms of idealised spectators, as opposed to normal, fallible (though properly-functioning) human beings.

<sup>17</sup> Two interpreters state something like the view I have been pressing (but don't do much more than that): Ingvar Johansson says that Hume "could very well have produced two definitions of every natural relation. Each such relation can be considered either in its associative and determining capacity or in its appearance as a mere comparison product" ("Hume's Scottish Kantianism", *Ruch Filozoficzny* LIX (2002): 421-53, 437). And David Owen, in the context of the two definitions, notes that a "philosophical relation is a matter of voluntarily comparing one idea with another" (*Hume's Reason* (Oxford: Oxford University Press, 1999), 152). Owen rejects the claim that the comparison of ideas can issue in causal judgement, however (152-3).

<sup>18</sup> Noonan, *Hume on Knowledge*, 151.

<sup>19</sup> Georges Dicker, *Hume's Epistemology and Metaphysics* (London: Routledge, 1998), 115.

<sup>20</sup> See §2 above; see also T. J. Richards, "Hume's two definitions of 'cause' ", *Philosophical Quarterly* 15 (1965): 247-53.

---

<sup>21</sup> Robinson, “Hume’s Two Definitions”, 166-7.

<sup>22</sup> Edward Craig, *The Mind of God and the Works of Man* (Oxford: Clarendon Press, 1987), 108.

<sup>23</sup> Simon Blackburn, “Hume and Thick Connexions”, in his *Essays in Quasi-Realism* (New York: Oxford University Press, 1993), 104.

<sup>24</sup> Angela Coventry, *Hume: A Guide for the Perplexed* (London: Continuum, 2007), 110.

<sup>25</sup> Actually Garrett argues that each definition can be given an “absolute” and a “subjective” reading, and if we keep the reading of both definitions the same, we get necessary extensional equivalence (*Cognition and Commitment*, 108-11). This is a neat trick. However, the passage just quoted forces a “subjective” reading – an account of when things “function psychologically” as cause and effect. The “absolute” reading of the first definition, according to which “constant conjunction” means “universal constant conjunction”, would be subject to the criticism raised in (a) above.

<sup>26</sup> I am not completely sure I have characterized Garrett’s view correctly; it may be that it belongs more comfortably in the “circumstances of belief formation” camp.

<sup>27</sup> John Wright, *The Sceptical Realism of David Hume* (Manchester: Manchester University Press, 1983), 130.

<sup>28</sup> *The Secret Connexion*, 208.

<sup>29</sup> Wright’s and Strawson’s discussions can be read as applying only to the *Enquiry* version of the two definitions, which omits reference to the natural/philosophical relation distinction, so it is perhaps not surprising (and not a glaring omission) that they do not mention it. On the other hand, given the similarity of the *Treatise* and *Enquiry*

---

versions of the definitions, it would be odd if the latter were aiming at something completely different to the former; see §5 below.

<sup>30</sup> See Buckle, *Hume's Enlightenment Tract*, 212-3, for a discussion of Hume's counterfactual rendition of the first definition. See Annette Baier, *Death and Character: Further Reflections on Hume* (Cambridge, MA: Harvard University Press, 2008), Chapter 10 for an account of why contiguity is missing. Hume also drops contiguity later in the *Treatise*, when he claims that there can be causation without contiguity in place in the cases where non-extended objects, which lack location, enter into causal relations (T 1.4.5; SBN 237).

<sup>31</sup> Galen Strawson, "David Hume: objects and power", in R. Read and K. Richman (eds), *The New Hume Debate* (London: Routledge, 2000), 31-2.