Abstract: This paper combines the ancient idea that causes necessitate their effects with Angelika Kratzer’s semantics of modality. On the resulting view, causal claims quantify over restricted domains of possible worlds determined by two contextually-determined parameters. I argue that this view can explain a number of otherwise puzzling features of the way we use and evaluate causal language, including the difference between causing an effect and being a cause of it, the sensitivity of causal judgements to normative facts, and the semantics of causal disagreements.

1. Introduction

If there is one constant in the history of philosophical investigation into the nature of causation, it’s the idea that causes necessitate their effects. Anscombe (1971) traces it back to Aristotle’s *Metaphysics*. Some early modern philosophers took it to require the inconceivability of causes without their effects,¹ a view which naturally leads to a version of occasionalism.² But Hume, despite agreeing that “[t]here is a NECESSARY CONNEXION to be taken into consideration” (Hume 1738, p. 77) between causes and their effects, famously looked elsewhere for the “source of our idea of power or necessity”, finding it in the “constant conjunction” of similar events and the “determination of the mind to pass from one object to its usual attendant” (Hume 1738, pp. 164-5).

Early critics of Hume’s complained that, on his view, “night is the cause of day, and day the cause of night. For no two things have more constantly followed each other since the
necessary connections in context

alex kaiserman

beginning of the world” (reid 1788, p. 249). to account for this, later writers distinguished merely accidently true generalizations from genuine laws of nature.³

this led to what we might call the ‘covering law’ condition on causation:⁴

(cl): c caused e only if c and e are events of type c and e, respectively, and it is a law of nature that events of type c are invariably followed by events of type e.⁵

not many philosophers of causation still believe that (cl) is true. for one thing, (cl) is inconsistent with the intuition many people claim to have that causation is possible in an indeterministic universe.⁶ but even assuming determinism, (cl) leads to what many see as a problematic proliferation of causes. suppose a short circuit occurs and a fire breaks out shortly afterwards.⁷ it’s consistent with the laws of nature that the short circuit occurs and the fire doesn’t, because there are no flammable materials nearby, or because there is no oxygen in the atmosphere, or because every atom in the andromeda galaxy simultaneously quantum-tunnelled through energy barriers 2.5 million years ago and the earth is destroyed by gravity waves before the fire has a chance to develop. indeed, “for any small region r of space at time t nothing much short of the state of the universe in a sphere with center r and whose radius is one light second (i.e. 186,000 miles) at t – 1 seconds is causally sufficient for determining what will occur...in r” (loewer 2007, p. 252). so according to (cl), what caused the fire was really a vast plurality of events, including the short circuit, but also the presence of oxygen, the absence of macroscopic quantum-tunnelling events,⁸ and plenty more besides.

mill embraced this conclusion. “the cause, then, philosophically speaking, is the sum total of the conditions [...] taken together; the whole of the contingencies of every description, which being realized, the consequent invariably follows” (mill 1843, p. 370). from among these things, mill insisted, “we have, philosophically speaking, no right to
give the name of cause to one of them, exclusively of the others” (Mill 1843, p. 366). But the problem is that we do routinely distinguish those events which caused an effect from those which were merely nomologically required for the effect to occur. In the case above, for example, it seems very natural in many contexts to describe the short circuit as the cause of the fire and the presence of oxygen as a mere ‘background condition’. As Hart and Honoré put it:

Mill’s doctrine [...] implies that every factor necessary for the occurrence of an event is equally entitled to be called ‘the cause’. Yet this is not the case: neither the plain man, nor the historian, uses the expression ‘cause’, or any related expression, in this way. For the contrast of cause with mere conditions is an inseparable feature of all causal thinking, and constitutes as much of the meaning of causal expressions as the implicit reference to generalizations does (Hart and Honoré 1985, p. 12).

Some millennia earlier, Plato expressed the same complaint:

If someone said that without bones and sinews and all such things, I should not be able to do what I decided, he would be right, but surely to say that they are the cause of what I do...is to speak very lazily and carelessly.

To call those things causes is too absurd [...] Imagine not being able to distinguish the real cause from that without which the cause would not be able to act as a cause (Plato, Phaedo 99a-b).

Some even interpret the proliferation of causes delivered by (CL) as a reductio of the very idea that causes necessitate their effects: “To treat causation as a type of necessity, or even as involving necessity, is to take causation as something it is not”, according to Mumford and Anjum (2011, p. 48). I think this is an overreaction. It’s well-known that modal claims in natural language rarely, if ever, quantify over every nomologically possible possible world. There are contexts in which I can truly utter the sentence ‘Eric can’t jump
eight metres’, even though there is a possible world in which he does jump eight metres (one in which he’s on the moon, for example); and there are contexts in which I can truly utter the sentence ‘Mary must be home by now’, even though there’s a possible world in which she isn’t home by now. The mechanisms by which the conversational context determines the restricted domains of possible worlds over which these sentences quantify is now fairly well understood. I propose to apply the insights of this literature to the idea that causes necessitate their effects. My hypothesis is that this approach can explain a number of otherwise puzzling features of the way we use and evaluate causal language, in everyday life, but also in science and the law.

The paper is structured as follows. Section 2 briefly introduces Kratzer’s semantics of modality. In section 3, I apply Kratzer’s framework to the idea that causes necessitate their effects. Sections 4-6 then use the resulting view to explain three neglected features of causal discourse: the distinction between causing an effect and being a cause of it, the influence of normative considerations on the causal claims we tend to make and endorse, and the semantics of certain kinds of causal disagreements. Section 7 compares my proposal with three alternative approaches, which appeal to pragmatics, ambiguity, and contrastivism respectively. Section 8 concludes.

2. Modality and Context

Modal claims are claims about what can, may, might, must, should, ought to, or has to be the case. It’s well-known that these claims can mean different things in different contexts. The modal auxiliary ‘must’, for example, has a deontic reading, as in ‘You must pay your taxes’, which has to do with rules or prescriptions; a bouletic reading, as in ‘You must apply for that job’, which has to do with an agent’s goals or desires; and an epistemic reading, as in ‘You must be tired’, which has to do with what an agent knows.
But there is also significant semantic variation even within one of these ‘flavours’ of modality. Suppose Karin commits murder. There’s an obvious sense in which Karin should be in jail in light of the law. But there’s also an obvious sense in which it should be the case, in light of the law, that Karin never committed murder in the first place, and hence isn’t in jail. What’s going on here?

The success of Kratzer’s (1981) solution is its ability to explain all this semantic variation without abandoning the assumption that modal auxiliaries express the same content in every context of use. The basic idea is that modal claims are relativized to a specification of what we’re ‘holding fixed’ and a method of ordering possible worlds. These together pick out a restricted domain of possible worlds – the ‘best’ worlds, on the relevant ordering, consistent with what we’re ‘holding fixed’. Something can or may or might be the case exactly if it’s the case in at least one of these worlds, and something must or should or has to or ought to be the case exactly if it is the case in every such world. One can generate different readings of modal claims by varying what we’re holding fixed or how we’re ordering possible worlds. ‘You must give to charity’ might be true on a moral ordering but false on a legal ordering, for example. If we hold fixed the fact that Karin committed murder, the legally ‘best’ worlds are all ones in which Karin is in jail. But if we hold nothing fixed, the legally ‘best’ worlds are all ones in which Karin doesn’t commit murder in the first place, and so isn’t in jail.

More precisely, modal auxiliaries on Kratzer’s view express functions with, not one, but three arguments – a proposition (e.g. the proposition that Karin is in jail) and two sets of propositions, the modal base $\mathbf{B}$ and the ordering source $\mathbf{O}$. Propositions are sets of possible worlds on Kratzer’s framework, so $\mathbf{B}$ is a set of sets of possible worlds. The set $\cap \mathbf{B}$, then, is the set of worlds in which every proposition in $\mathbf{B}$ is true. $\mathbf{O}$ then defines a partial order $\leq_0$ over $\cap \mathbf{B}$ as follows: for all $w, u$ in $\cap \mathbf{B}$, $w \leq_0 u$ if and only if whenever an
element of $O$ is true in $w$, it’s also true in $u$. Call the set of maximal elements of $\cap B$ on $\leq_0$, ‘$\text{MAX}_{\leq_0}B$’. Then ‘It must be the case that $\varphi$’ is true if and only if $\varphi$ is true in every element of $\text{MAX}_{\leq_0}B$; ‘It may be the case that $\varphi$’ is true if and only if $\varphi$ is true in some element of $\text{MAX}_{\leq_0}B$; and similarly for other modal locutions.’

There are different kinds of modal base – sometimes we’re interested in what is possible holding fixed all and only those things we know, sometimes what is possible holding fixed certain salient facts about the circumstances. There are different kinds of ordering source – sometimes we’re interested in those worlds which satisfy the most legal, moral, stereotypical or social ideals, sometimes those worlds which satisfy the most goals or desires or teleological ends. We can gesture at these parameters explicitly – *given* that Karin has committed murder, then *legally speaking*, she should be in jail. But in most cases we needn’t bother, since the conversational context often succeeds in filling in the extra argument places for us. “When we talk to each other, we hardly ever make explicit in view of which circumstances something should be necessary or possible. We may give hints. Usually people understand. And they all understand in pretty much the same way” (Kratzer 1981, pp. 53-4).

Kratzer’s theory leaves many questions unanswered and it’s not entirely uncontroversial. But it represents the closest we have to orthodoxy about the semantics of modality. It has been enthusiastically applied to a wide variety of expressions, from conditionals (Kratzer 2012) to generics (Krifka et al. 1995). But there have been few attempts to apply it to the semantics of causal claims. This is a puzzling lacuna. Although many philosophers have found it plausible that there are deep conceptual connections between causation and modality, philosophical analyses of everyday causal claims have tended to ignore all the rich semantic complexity of modality in natural language; complexity which carries over into the semantics of causal claims too.
3. Causation and Context

I propose to apply Kratzer’s theory to causal claims. On the view I will defend, ‘cause’ expresses a function (the same function in every context) with, not two, but four arguments: A plurality of events $X_1, \ldots, X_n$ (the causes), an event $Y$ (the effect), a modal base $B$ and an ordering source $O$. As with modal claims, the conversational context is often sufficient to fix $B$ and $O$ without them having to be explicitly specified (though we can clarify, for example, that the short circuit caused the fire given that there was oxygen in the atmosphere).

Now here’s a plausible gloss on the idea that causes necessitate their effects:

NECESSARY CONNECTIONS (NC): $X_1, \ldots, X_n$ collectively caused $Y$ only if, given that $X_1, \ldots, X_n$ all occurred, $Y$ had to occur.

In sentences of the form ‘Given that $\phi$, it had to be that $\psi$’, the role of the first clause is to add the proposition expressed by $\phi$ to the modal base relative to which the second clause is evaluated. In particular, if $x_1, \ldots, x_n$ are the propositions that $X_1, \ldots, X_n$, respectively, occurred, the sentence ‘Given that $X_1, \ldots, X_n$ all occurred, $Y$ had to occur’ is true relative to a modal base $B$ and ordering source $O$ if and only if ‘$Y$ had to occur’ is true relative to $B \cup \{x_1, \ldots, x_n\}$ and $O$. Properly spelled out, therefore, (NC) reads as follows:

(NC*): $X_1, \ldots, X_n$ collectively caused $Y$ relative to $B$ and $O$ only if $Y$ occurs in every element of $\text{MAX}_{\text{co}} \cap (B \cup \{x_1, \ldots, x_n\})$.

To determine whether the short circuit caused the fire relative to $B$ and $O$, for example, we take the set of possible worlds in which all the elements of $B$ are true, throw out all those worlds in which the short circuit doesn’t occur, order the elements that remain
using O, and check whether the fire occurs in every top-ranked world of this ordering – if it doesn’t, the short circuit didn’t cause the fire relative to B and O.

I’ve introduced a hypothesis about the logical forms of causal claims – that they contain two parameters usually determined by the conversational context – and reinterpreted the idea that causes necessitate their effects in light of this hypothesis. Since (NC*) is just a necessary condition on causation, I don’t purport to have provided an analysis of causal claims. Nevertheless, (NC*) can explain a number of otherwise puzzling semantic phenomena (sections 4-6) better than the alternatives (section 7). Or so I will argue.

4. Causes and Causing

Consider the following sentence:

(1) Jaya and Fatima lifted the table.

(1) is ambiguous. Read distributively, it follows from (1) that Jaya lifted the table and Fatima also lifted the table (at a different time). But the more natural reading is the collective one, according to which Jaya and Fatima lifted the table together. These are distinct states of affairs. On the distributive reading the table is lifted twice; on the collective reading it is only lifted once, even though it is lifted by two people. Lifting relates pluralities to individuals, in general; and a plurality of people can collectively lift a table without any one of the plurality individually lifting it.

Exactly the same is true of (2):

(2) The driver’s drunkenness and the rainstorm caused the car crash.

(2) is also ambiguous. Read distributively, it follows from (2) that the drunkenness caused the crash and the rainstorm also caused the crash. On the more natural collective
reading, however, the drunkenness and the rainstorm \textit{jointly} caused the crash. These are distinct states of affairs. On the distributive reading, the crash was caused twice over – it was \textit{overdetermined}, to put it another way\textsuperscript{14} – whereas on the collective reading the crash was only caused once, even though it was caused by two events. \textit{Causing} (and, for that matter, \textit{opening}, \textit{authoring}, \textit{preventing}, and so on) relates pluralities to individuals, in general; and a plurality of events can collectively cause an effect without any one of the plurality individually causing it.

The ambiguity of (2) is naturally explained by (NC*) – on the distributive reading, the drunkenness and the rainstorm \textit{individually} necessitated the crash, whereas on the collective reading they only \textit{jointly} necessitated it. More precisely, if $d$ is the proposition that the drunkenness occurred and $r$ is the proposition that the rainstorm occurred, (2) is true on its distributive reading relative to $B$ and $O$ only if the crash occurs in every element of $\text{MAX}_{\text{so}} \cap (B \cup \{d\})$ \textit{and} every element of $\text{MAX}_{\text{co}} \cap (B \cup \{r\})$; whereas on its collective reading, (2) is true relative to $B$ and $O$ only if the crash occurs in every element of $\text{MAX}_{\text{co}} \cap (B \cup \{d, r\})$.

If a book is collectively authored by multiple people, each of those people was \textit{an author} of the book. To be an author of a book, in other words, is to be one of the people who collectively authored it. The same is true for causation. If an effect was collectively caused by a plurality of events, each of those events was \textit{a cause} of the effect. To be a cause of an effect, in other words, is just to be one of the events that collectively caused it. ‘$X$ is a cause of $Y$’ and ‘$X$ caused $Y$’ are therefore \textit{not synonymous}, notwithstanding a widespread tendency among philosophers to use them interchangeably. If $X$ caused $Y$, then $X$ was trivially a cause of $Y$. But the converse doesn’t follow. On the collective reading of (2), the rainstorm was \textit{a cause} of the crash, but it didn’t \textit{cause} it – what caused it was the drunkenness and the rainstorm taken together.
The distinction between *causing* and being *a cause* of an effect is context-sensitive, on my view. Suppose, by analogy, that Eric claims that he can jump eight metres. Consider the following exchange between Yena the athlete and David the facetious physicist:

Yena: Eric can’t jump eight metres!
David: Well technically speaking he *can* – anyone can jump eight metres on the moon, for example.

David’s utterance is literally true, of course. But intuitively, it doesn’t contradict Yena’s utterance because David has changed the context (this is the sense in which David is being facetious). Yena’s utterance was true (and David’s false) in the context in which she made it, because, in particular, in no possible world consistent with the modal base in that context is Eric on the moon – we’re ‘holding fixed’ the fact that Eric is on earth. But everyday utterances are governed by what Lewis calls ‘a rule of accommodation’: “what you say makes itself true, if at all possible, by creating a context that selects the relevant features so as to make it true” (Lewis 1986c, p. 251). David’s utterance succeeds in changing the context to one relative to which it expresses a truth, namely one which quantifies over a greater number of possible worlds, at least one of which is a world where Eric jumps eight metres on the moon.

Now imagine a similar exchange between David and Ulla the engineer:

Ulla: The short circuit caused the fire.
David: Well technically speaking the short circuit was only one cause among many – the presence of oxygen was also a cause of the fire, for example.

Again, what David says is perfectly true, but it changes the context. Ulla’s utterance was true in the context in which she made it, because, in particular, in no possible world consistent with the modal base in that context does there fail to be oxygen in the
atmosphere – we’re ‘holding fixed’ the presence of oxygen. But David’s utterance succeeds in changing the context to one which quantifies over more possible worlds, at least one of which is a world where the short circuit occurs without any oxygen in the atmosphere, and hence without the fire. Relative to this context, only the short circuit and the presence of oxygen taken together necessitated the fire, so that the short circuit was a cause of the fire, but didn’t cause it.\(^{15}\)

Lewis famously claimed to be “concerned with the prior question of what it is to be one of the causes (unselectively speaking)” of an effect, and his analysis was “meant to capture a broad and nondiscriminatory concept of causation” (Lewis 1986a, p. 162). One way of interpreting this comment is to take Lewis to be giving an account of what it is to be a cause of an effect, relative to every nomologically possible world; relative, that is, to an empty ordering source and a modal base that only includes the laws of nature. Relative to these parameters, the fire was collectively caused by the short circuit, the presence of oxygen, the absence of macroscopic quantum-tunnelling events, and so on, so that each was a cause of the fire, though none of them (individually) caused it. Analyzing what it is to be a cause of an effect in this sense is a perfectly respectable metaphysical project, analogous to analyzing what it is for something to be metaphysically possible (i.e. possible, unrestrictedly speaking). But it’s not the whole story, at least if what we want is a semantic theory of causal claims – a theory of what causal claims (in contexts) mean. The fact is that we do routinely – and, it seems, truly – make claims about what caused an effect, such as ‘The short circuit caused the fire’, that fail to mention even a fraction of the vast plurality of events that collectively caused the effect in this unrestricted sense. On the view presented here, this isn’t down to “the inaccuracy of common discourse” (Mill 1843, p. 370), but rather down to the antecedently accepted and well-understood mechanism by which the conversational
context restricts the domains of possible worlds over which certain kinds of expression quantify.

5. Causes and Norms

Consider the following case, from Clarke et al. (2015):

(Collision)

Two cars, one driven by Greta and the other driven by Rachel, were approaching an intersection. Greta had a green light. Rachel had a red light, but she wasn’t paying attention. The lights stayed that way. Neither driver stopped, and their cars collided (Clarke et al. 2015, p. 282).

Now consider the following claims:

(3) Rachel’s driving into the intersection caused the collision.

(4) Greta’s driving into the intersection caused the collision.

On average, experimental subjects presented with Collision agreed with (3) and disagreed with (4), despite the fact that the collision wouldn’t have occurred if either Greta hadn’t driven into the intersection or Rachel hadn’t driven into the intersection. The only relevant difference between the two events seems to be a normative one: Driving through green lights is what one is supposed to do, whereas one isn’t supposed to drive through red lights. These judgements have been widely replicated in a variety of different settings and appear to be remarkably robust.

(NC*) can explain what’s going on here. The asymmetry in our causal judgements about Collision is due to the fact that the details of the case encourage us to evaluate (3) and (4) relative to something like the following ordering source:
\[O = \{\text{Everyone approaching a red light stops, Everyone approaching a green light carries on}\}.\]

Kratzer calls this a ‘stereotypical’ ordering source – red lights and green lights are associated with certain stereotypical states of affairs, namely stopping and continuing respectively.

Let \(g\) be the proposition that Greta continues into the intersection and let \(r\) be the proposition that Rachel continues into the intersection. Given a suitable modal base \(B\), the elements of \(\text{MAX}_{\omega_0} \cap B\) – the ‘top-ranked’ worlds of this ordering – are all ones in which Rachel stops at the intersection, Greta carries on, and the collision doesn’t occur. The elements of \(\text{MAX}_{\omega_0} \cap (B \cup \{g\})\) are therefore also ones in which Rachel stops at the intersection, Greta carries on, and the collision doesn’t occur. Hence it’s inconsistent with \((\text{NC}^*)\) that Greta’s continuing into the intersection caused the collision, relative to \(B\) and \(O\) – the collision didn’t have to occur, even given that Greta’s continuing into the intersection occurred. The elements of \(\text{MAX}_{\omega_0} \cap (B \cup \{r\})\), by contrast, are ones in which Rachel continues into the intersection, as does Greta, and so the collision does occur. Assuming the other necessary conditions for causation are satisfied, \((\text{NC}^*)\) implies that Rachel’s continuing into the intersection caused the collision, relative to \(B\) and \(O\) – given that her continuing into the intersection occurred, the collision had to occur, in the relevant sense. Hence in this context, \((3)\) is true and \((4)\) is false. That’s why subjects agree with \((3)\) and disagree with \((4)\).

Of course, we can consider the same case in different contexts. Relative to an empty ordering source (and the same modal base), for example, neither Greta’s nor Rachel’s driving into the intersection individually caused the collision. Rather, the collision was collectively caused by both events, so that Greta’s driving into the intersection was a
cause of the collision, as was Rachel’s. (NC*) therefore predicts that we will elicit different responses from subjects if we deliberately undermine the salience of the stereotypical ordering source suggested by the case. For example, Clarke et al. found that when subjects were instructed to think about what caused the collision “considering the physics of the situation” (Clarke et al. 2015, p. 288), they were much more likely to agree that both events were causes of the collision.18

The phenomenon exemplified by Collision is pervasive in causal discourse. Causal claims are routinely evaluated relative to many different kinds of ordering source, in everyday life, but also in science and the law. Consider, for example, the legal case of McKew v. Holland,19 in which an employee of the defendant company sustained an injury in a workplace accident. As a result, his left leg was weakened and would frequently become numb. Three weeks after the accident, the plaintiff lost his balance while descending a steep staircase without a handrail. He fell and sustained a further injury. Neither injury would have occurred but for the company’s negligence. But although the company accepted liability for the first injury, it appealed against the decision to hold it liable for the second, on the grounds that its negligence didn’t cause the second injury. The House of Lords agreed:

If a man is injured in such a way that his leg may give way at any moment he must act reasonably and carefully. It is quite possible that in spite of all reasonable care his leg may give way in circumstances such that as a result he sustains further injury. Then that second injury was caused by his disability which in turn was caused by the defendant’s fault. But if the injured man acts unreasonably […] what follows must be regarded as caused by his own conduct and not the defendant’s fault.20

It’s difficult to make any sense of this reasoning on standard philosophical theories of causation. How could the causal history of the second injury have anything to do with
whether the plaintiff’s conduct was ‘reasonable’? (NC*) can provide an explanation. The causal claims in this passage are being evaluated relative to a bouletic ordering source. According to (NC*), the negligence caused the injury in this context only if the plaintiff had to go down the stairs, in the sense of there being no reasonable alternative to his doing so in light of his goals and desires. If there was a reasonable alternative, then the negligence didn’t cause the second injury, since even given that the negligence occurred, the second injury didn’t have to occur.

Bouletic uses of modal terms appear frequently in causal reasoning in the law. In a similar case wherein the plaintiff, fearing for his safety, jumped from the defendant’s coach and broke his leg, Lord Ellenborough argued that the defendant’s negligence caused the injury only if the negligence placed the plaintiff “in such a situation that he must adopt a perilous alternative”, thereby creating “a necessity for what he did”. The defendant’s failure to fix a defective train door in Adams v. The Lancashire and Yorkshire Railway Company was found not to have caused the plaintiff’s injury sustained while trying to close it, since the plaintiff didn’t have to try so hard to shut the door: “[I]t was not even necessary to do so to avoid inconvenience”.

We can find evidence of the same phenomenon in scientific causal claims too. Here is a well-known example from molecular biology. The output of the process of protein synthesis in cells counterfactually depends, not only on the specificities of the cell’s DNA, but also on the specificities of a whole host of other molecules, including tRNA molecules, the carriers of amino acids. But although “everyone knows that many different molecules and cellular structures play necessary roles in the in vivo syntheses” of proteins, “biologists and philosophers often talk as if DNA ‘produces’” them, with the other molecules merely mediating this causal influence (Waters 2007, p. 553). Again, these distinctions are difficult to explain on standard philosophical theories of causation.
But (NC*) can explain them, on the assumption that biologists are often interested in what caused an effect relative to teleological ordering sources. As Weber points out:

Substituting tRNAs [...] for molecules with different specificities [...] is not consistent with continuing biological functioning of the process of protein synthesis. For if the cell suddenly contains different tRNAs [...] this will affect the sequence of other protein molecules made by the cell, which makes it impossible for it to survive (Weber forthcoming).

In other words, the cell has to have the particular tRNA molecules it in fact has to ensure continued proper functioning. A certain DNA sequence therefore necessitates the production of a certain kind of protein, relative to those teleologically top-ranked possible worlds in which the cell is functioning properly. By contrast, a cell doesn’t have to have the particular DNA it in fact has to ensure continued proper functioning – Gibson et al. (2010) recently succeeded in replacing a bacterium’s entire genome with one created in the laboratory, after which the cell remained capable of continuous self-replication. Particular tRNA molecules alone do not therefore necessitate the production of a certain kind of protein, even relative to those possible worlds in which the cell is functioning properly, since in some of these worlds the cell contains different DNA and therefore produces a different protein.

6. Causal Disagreement as Normative Disagreement

The view I have defended can be further illustrated by how it deals with a curious species of causal disagreement. Gradable adjectives – words like ‘tall’, ‘flat’, ‘rich’, and so on – are generally thought to express relations between objects and comparison sets (e.g. Kennedy and McNally 2005). For example, an utterance of ‘Kim is tall’ expresses the proposition that Kim is tall for X, where the value of X is a set of individuals determined by the context. This means that there are two different ways of disagreeing about the
truth of the proposition expressed by ‘Kim is tall’ in a particular context. Suppose we’re in a context where ‘Kim is tall’ expresses the proposition that Kim is tall for a basketball player. I think ‘Kim is tall’ is true in such a context, whereas you think it’s false. This might be because we disagree about Kim’s height. But it might also be because we disagree about the average height of basketball players, and hence about how tall Kim has to be in order to be tall for a basketball player.

We get exactly the same phenomenon with causal claims. Consider the following case:

**Bankruptcy**

A company is made up of low-paid workers and well-paid managers. The managers decide that they deserve more money for the work that they do and give themselves a large pay-rise; at the same time, the workers organise an unannounced strike over pay and conditions. Neither group knew of the other’s plans. The company subsequently falls into financial difficulties and declares bankruptcy.

Consider a disagreement between a worker and a manager. Let’s suppose that they agree on all the counterfactual facts – they agree, in particular, that the company wouldn’t have had to declare bankruptcy if the strike had occurred without the pay-rise or if the pay-rise had occurred without the strike. Nevertheless, the manager makes utterances like “It wasn’t the pay-rise that caused the bankruptcy, it was the strike” and the worker replies with utterances like “It wasn’t the strike that caused the bankruptcy, it was the pay-rise”. The source of this disagreement seems to be a *normative* disagreement – the manager believes that the pay-rise was morally required, given the work she does, and that the strike was unnecessary (the workers could have resolved their grievances in other ways); whereas the worker believes that the strike was morally required, given the working conditions and managerial intransigence, and that the pay-rise was unnecessary (the managers could have done without the extra money).
These kinds of causal disagreements are pervasive in political, historical, legal and everyday discourse. But again, they're difficult to explain on standard philosophical theories of causation. How could the manager and the worker agree on all the counterfactual facts and yet fail to agree on the causal facts? The most common diagnosis is that these disagreements aren’t really disagreements about causation at all; rather the worker and the manager really agree on all the causal facts, but disagree about something else (such as who was morally responsible for the bankruptcy), and mistakenly phrase their disagreement in causal terms.\(^{27}\)

On my view, we can and should take this disagreement at face value. The manager and the worker do disagree on the causal facts; and they do so precisely because they disagree on the moral facts. Although the manager and the worker agree that they’re in a context where causal claims are being evaluated relative to a moral ordering source, they disagree about what the moral ordering source is. Hence, they disagree about which sets of possible worlds their causal claims are quantifying over. Relative to a suitable modal base, the worker thinks that every top-ranked world on the moral ordering source is one in which the strike occurs, but not every top-ranked world is one in which the pay-rise occurs. Hence she believes that, given that the pay-rise occurred, the bankruptcy had to occur, although the same is not true of the strike. On the other hand, the manager thinks that every top-ranked world on the moral ordering source is one in which the pay-rise occurs, but not every top-ranked world is one in which the strike occurs. Hence she believes that, given that the strike occurred, the bankruptcy had to occur, although the same is not true of the pay-rise.

Again, this kind of disagreement arises only in contexts where a particular ordering source is salient. The manager and the worker would presumably agree that, physically speaking (i.e. relative to an empty ordering source), both the pay-rise and the strike were
causes of the bankruptcy. But that’s not what’s at issue. What’s at issue is what caused the bankruptcy \textit{morally} speaking. And they disagree about that, precisely because they disagree about the moral facts.\footnote{28}

7. Alternatives

I’ve argued that we can explain a number of features of causal discourse by applying Kratzer’s semantics of modality to (NC). Many alternative explanations of these phenomena have been offered in the literature. I don’t have space to consider them all, but it’s worth comparing my view with three popular alternative proposals – the \textit{pragmatic} response, the \textit{ambiguity} response and the \textit{contrastivist} response.

Recall that subjects presented with \textit{Collision} tend to \textit{agree} with (3) and \textit{disagree} with (4):

(3) Rachel’s driving into the intersection caused the collision.

(4) Greta’s driving into the intersection caused the collision.

On my view, this is simply because, in the relevant context, (3) is true and (4) is false. But the proponent of the \textit{pragmatic} response disagrees. In fact, she claims, (3) and (4) are both true (in every context). But an assertion of (4), in the relevant context, would be \textit{pragmatically inappropriate}; and that’s why subjects are reluctant to endorse it. As Lewis puts it, “[t]here are ever so many reasons why it may be inappropriate to say something true. It might be irrelevant to the conversation, it might convey a false hint, it might be known already to all concerned...” (Lewis 2000, p. 196).

As others have pointed out, however (see McGrath 2005; Schaffer 2013), the pragmatic response doesn’t really explain the data. We aren’t merely unwilling to assert (4); we’re
also more than happy to assert its negation. The following seems like a perfectly natural thing to say in the context:

(5) Greta’s driving into the intersection didn’t cause the collision. It was Rachel’s driving into the intersection which caused it; Greta’s driving into the intersection was just a background condition.

It doesn’t follow from the fact that it’s pragmatically inappropriate to utter a true sentence that it would be appropriate to utter its negation. If you ask me for directions and I respond with ‘The earth is round’, my utterance is pragmatically inappropriate, since the information I supply is irrelevant. But that doesn’t mean that it would have been acceptable for me to respond instead with ‘The earth isn’t round’. So simply pointing out that it would be pragmatically inappropriate to utter (4) in some contexts doesn’t explain why it is appropriate in those contexts to utter (5).

There are admittedly some instances of so-called metalinguistic negation in which the word ‘not’ is used to register dissatisfaction with the pragmatic propriety of a would-be utterance rather than with its literal falsity. Consider the following sentence:

(6) The water isn’t hot; it’s scalding!

(6) is contradictory if the (contracted) ‘not’ is interpreted as expressing logical negation – if something is scalding, then necessarily it’s hot. But there is nevertheless a consistent reading of (6), one where the ‘not’ is interpreted metalinguistically. On this reading, (6) says that an utterance of ‘It’s hot’ was or would have been inappropriate, because it would have pragmatically implicated the false proposition that the water isn’t scalding. One might therefore be tempted to say the same about (5). It has a true reading, one where the ‘not’ is interpreted metalinguistically. Perhaps, for example, (5) says that an utterance of ‘Greta’s driving into the intersection caused the collision’ would have been
inappropriate, because it would have pragmatically implicated the false proposition that
Greta was to blame for the collision.

The only problem with this proposal is that there isn’t any actual evidence that ‘not’ is
being used metalinguistically in (5). Metalinguistic readings are fickle (see Horn 1989).
Firstly, they rely on specific intonation contours. Placing the stress in a different place in
(6) makes the metalinguistic reading much harder to access:

(7) # The water isn’t hot. It’s scalding.

Secondly, metalinguistic readings disappear in ‘concessive’ constructions like (8):

(8) # The water isn’t hot, although it is scalding.

Although (6) has a clear true reading, (7) and (8) strike us as odd or even contradictory.
But we don’t observe the same patterns for (5):

(9) Greta’s driving into the intersection didn’t cause the collision. It was just a
background condition.
(10) Greta’s driving into the intersection didn’t cause the collision, although it was a
background condition.

Speaking for myself, I find (9) and (10) to be just as acceptable as (5). This suggests that
the ‘not’ in (5) is not being interpreted metalinguistically. These tests are not decisive, of
course, but they work well enough to make us suspicious of unqualified appeals to
pragmatics.

The second response, the ambiguity response, postulates an ambiguity in the word
‘cause’. This view “treats our ordinary talk of causation as shifting between two distinct
‘concepts’ of cause”, resulting in “an ambiguity that could be resolved if each concept was
given a different name” (Godfrey-Smith 2009, p. 327). One is the “broad and non-
discriminatory notion of causation” (Lewis 1986a, p. 162) and the other is a narrower “selective notion of causation” (Broadbent 2012, p. 463). We might call the former ‘causal influence’ and the latter ‘explanatory relevance’ (Strevens 2008). Either way, the idea is that Greta’s driving into the intersection ‘caused’ the collision in the broad but not the narrow sense, whereas Rachel’s driving into the intersection ‘caused’ the collision in both the broad and the narrow sense. (5) is an acceptable thing to say because it has a true interpretation, one where ‘cause’ is interpreted narrowly.

Again, the problem with this view is that there is no actual evidence that ‘cause’ is ambiguous in this way. An ambiguity between a broad and a narrow understanding of a word is known as a privative ambiguity. Zwicky and Sadock (1975) suggest a number of tests for determining whether a term is privatively ambiguous. One of these is to check whether it is “possible, without contradiction, to assert the general while denying the specific” (Zwicky and Sadock 1975, p. 7). Compare the following two sentences:

(11) The aeroplane climbed, but it didn’t climb.

(12) # The aeroplane crashed, but it didn’t crash.

There is no way of understanding (12) on which it isn’t a contradiction. But there is a way of understanding (11) on which it isn’t a contradiction. This suggests that ‘climb’ is ambiguous between a narrow literal meaning (where climbing requires arms and legs) and a broader more metaphorical meaning.

Now consider the following sentence:

(13) Greta’s driving into the intersection caused the collision, but it didn’t cause the collision.
Speaking for myself, I find it hard to understand (13) as anything other than a straightforward contradiction. This suggests that ‘cause’ is not ambiguous. Again, this test is not decisive, but it works well enough to make us suspicious of unqualified appeals to ambiguity.

The last response I’ll consider is the constrastivist response (Schaffer 2005, 2013; Hitchcock 2007). According to contrastivism, causation is a four-place, contrastive relation – roughly speaking, X rather than X* caused Y rather than Y* if and only if, had X* occurred instead of X, Y* would have occurred instead of Y. Sentences of the form ‘X caused Y’ express contrastive propositions, on this view, where the contrast events are determined by the conversational context. The idea is that ‘’deviant’ events tend to leap out as especially salient to people and tend to trigger thoughts of the more normal alternative, while ‘default’ events tend to duck out of view and not trigger thoughts about alternatives at all” (Blanchard and Schaffer forthcoming). For example, Rachel’s driving into the intersection is a ‘deviant’ event, since she had a red light. It triggers thoughts of a more ‘normal’ alternative, namely Rachel’s stopping at the intersection. Hence an utterance of (3) is naturally interpreted as expressing the (true) proposition that Rachel’s driving into the intersection rather than stopping caused there to be a collision rather than no collision. But Greta’s driving into the intersection was a ‘default’ event, since she had a green light. Hence it doesn’t trigger thoughts of any alternatives. It follows that an utterance of (4) cannot be interpreted as expressing a contrastive proposition, “nor is it obvious what if any interpretation it should receive” (Schaffer 2013, p. 52).

I think the contrastivist is wrong about the data here. If there is indeed no salient alternative to Greta’s driving into the intersection, (4) is uninterpretable on the contrastive account, since one of the arguments of the function expressed by ‘cause’ remains unspecified by the context. But if (4) is uninterpretable then so is its negation,
for the same reason. Hence (5) is also uninterpretable according to the contrastivist – it’s neither true nor false, since it fails to express a proposition at all. This seems like the wrong result. Speaking for myself, (5) just seems straightforwardly true. Schaffer himself seems to acknowledge as much in a footnote (he is concerned with a different sentence, but the same considerations apply):

\textit{Lacuna:} if [4] does not receive any natural interpretation than [sic] its denial should not either, which does not quite fit that data [...] So it would be smoother for me to say that [4] does receive some interpretation as a contrastive falsehood [...] But I do not currently have any contrastive falsehood to suggest for the role (Schaffer 2013, pp. 61-2).

These considerations aren’t decisive, but again, they should make us suspicious of unqualified appeals to contrastivism to explain these phenomena.

I’ve considered three alternative explanations of our judgements concerning (3), (4) and (5). Though I don’t purport to have provided any knock-down arguments, I’ve shown that there are serious problems with each of them. This, at the very least, is reason to take the view I have defended seriously.

\textbf{8. Conclusion}

Mackie famously argued that “causal statements are commonly made in some context, against a background which includes the assumption of some causal field” (Mackie 1974, p. 34). The ‘causal field’, for Mackie, is not “part of a cause, but is rather a background against which the causing goes on” (Mackie 1974, p. 63). The view defended in this paper can be seen as a contemporary development of Mackie’s insight, obtained by combining the ancient idea that causes necessitate their effects with a familiar framework borrowed from Kratzer’s theory of modality.
You might agree with everything I’ve said so far. You might, that is, agree that causal utterances often are evaluated relative to deontic, bouletic and teleological ordering sources. Nevertheless, you might also insist that they shouldn’t be evaluated in this way, at least for certain purposes. Stapleton accepts that “causation’ is a term we use to express diverse information about the world”, but suggests that in legal contexts, “clarity is promoted if we use the term ‘causation’ to refer to the information yielded by only one type of inquiry” (Stapleton 2008, p. 432) one which is “untainted by normative interrogations and controversies” (Stapleton 2008, p. 446). In a similar vein, some philosophers have argued that causal explanations in biology “must take the whole developmental matrix into consideration”, rather than focusing on DNA, a position known as the ‘causal parity thesis’ (Faye 2014, p. 152; see also Oyama et al. 2001). The question of whether such stipulated constraints on the context-sensitivity of causal language have their practical merits is beyond the scope of semantics, and therefore beyond the scope of this paper. It’s worth pointing out that biology and the law have arguably managed just fine without them, however. Indeed it’s worth remembering why natural languages contain context-sensitive elements in the first place: They allow us to exploit rich and subtle features of the circumstances in which we find ourselves to successfully communicate a large variety of different kinds of content with minimal linguistic resources. What’s not to like that about that?

Notes

1 “[A]n entire cause, is the aggregate of all the accidents both of the agents how many soever they be, and of the patient, put together; which when they are all supposed to be present, it cannot be understood but that the effect is produced at the same instant” (Hobbes 1655, pp. 121-2; my emphasis).
“A true cause as I understand it is one such that the mind perceives a necessary connection between it and its effect. Now the mind perceives a necessary connection only between the will of an infinitely perfect being and its effects. Therefore, it is only God who is the true cause and who truly has the power to move bodies” (Malebranche 1674, p. 450).

For example: “There are sequences, as uniform in past experience as any others whatever, which yet we do not regard as cases of causation, but as conjunctions in some sort accidental. Such, to an accurate thinker, is that of day and night […] We may define, therefore, the cause of a phenomenon, to be the antecedent, or the concurrence of antecedents, on which it is invariably and unconditionally consequent” (Mill 1843, p. 377).

Davidson (1980, p. 208) calls this the “Principle of the Nomological Character of Causality”.

Note that (CL) is just a necessary condition on causation. Those seeking an analysis have typically combined (CL) with other necessary conditions. According to Hume, for example, a cause must also occur prior to, and spatially contiguous with, its effect (Hume 1738, pp. 75-6).

For example, see Mackie (1974, pp. 40-3). I won’t address the issue of indeterministic laws here. It’s worth pointing out, however, that both of our two best candidates for a fundamental physical theory, general relativity and quantum mechanics, are perfectly deterministic. Proponents of the so-called ‘GRW theory’ have responded to perceived conceptual difficulties with quantum mechanics by postulating an additional indeterministic dynamical process, called ‘collapse’; but there are other deterministic alternatives (e.g. De-Broglie-Bohm theory) as well as interpretations of unitary quantum mechanics that take it seriously on its own terms (e.g. the Everett or ‘many worlds’ interpretation).

This example is adapted from one in Mackie (1965).

I wish to remain neutral here on the question of whether absence-talk picks out some sui generis entities or merely events under negative descriptions. See Lewis (1986a) and Schaffer (2005).

I’m assuming here that $\text{MAX}_{\mathcal{A}\cap\mathcal{B}}$ is always non-empty. This is called the limit assumption, and it’s endorsed by Stalnaker (1981) and rejected by Lewis (1973) and Kratzer. If the limit
assumption is false we need to massage the definitions above to avoid trivial results (see Kratzer 1981). I’ll ignore these complications from now on for simplicity.

10 In particular, it cannot by itself explain why some modals have preferred readings (e.g. ‘might’ likes to be epistemic, ‘can’ likes to be circumstantial), as well as various phenomena involving the interaction of modality with tense and aspect (Hacquard 2009).

11 Some think that utterances of sentences like ‘Karin should be in jail’ don’t express anything truth-evaluable, but rather express (in a different sense of ‘express’) something like the utterer’s disapproval of Karin’s actions (Horgan and Timmons 2006). Others think that sentences like ‘Karin should be in jail’ express the same proposition in every context of use, but these propositions are themselves true or false only relative to a context of assessment (MacFarlane 2005). Yet others distinguish what is said by an utterance from what is semantically expressed by it, and argue that indefinitely many propositions are said by an utterance of ‘Karin should be in jail’, even though only a single proposition is semantically expressed by it (Cappelen and Lepore 2005). I mention these alternatives merely to set them aside, but it’s worth thinking about how my view could be adapted to fit these various approaches.

12 The only exception I’m aware of is Szabó and Knobe (2013), who seek to explain certain patterns of empirical data by assuming that participants evaluate target statements by replacing them with “modal proxies”. My view is in the spirit of Szabó and Knobe’s project, although it goes further in that it actually incorporates Kratzer’s framework into the semantics of causal claims, and applies the resulting theory to a much broader range of phenomena.

13 When I say that Jaya and Fatima lifted the table together, I don’t mean that what lifted the table was the set containing Jaya and Fatima, or the mereological fusion of Jaya and Fatima, or indeed any other single thing. In other words, I believe that there is such a thing as irreducibly plural predication. See McKay (2006).

14 Unger (1977) argues that overdetermination in this sense is incoherent, since he thinks that the distributive readings of sentences like (2) are contradictory. I don’t endorse that view, though Unger mounts a convincing case.
An anonymous referee points out that one can felicitously say things like ‘The short circuit caused the fire, but of course it didn’t cause it alone’. One might conclude from this that ‘\(X\) caused \(Y\)’ and ‘\(X\) was a cause of \(Y\)’ are synonymous, although the former carries a cancellable presupposition that \(X\) was the only cause of \(Y\). But another explanation is that the context changes mid-sentence: The first clause is evaluated relative to a restricted set of possibilities and the ‘but of course’ indicates a change in context to a wider set of possibilities, relative to which the short circuit wasn’t the only cause. One finds similar effects in modal claims: On the natural reading of ‘John must pay his debts, but of course he can’t’, the first clause is evaluated relative to one set of possibilities and the second clause relative to a different set of possibilities.

See the results of Clarke et al.’s (2015) ‘First experiment’.

Clarke et al. re-ran their experiment with four different target statements and four different experimental primes, each specifically designed to urge the participants to respect the counterfactual symmetry of the case in their causal assessments. “No matter how plain we made it […] that what Greta did made a difference to whether the outcome occurred, participants tended to take Rachel but not Greta to be one of the causes” (Clarke et al. 2015, p. 283). See also Hitchcock and Knobe (2009), Knobe (2010), and references therein.

“[F]ocusing participants’ attention in a certain way – emphasizing the physics of a situation – can increase the frequency with which they give egalitarian responses” (Clarke et al. 2015, p. 289). This was the only experimental prime that had any significant effect on responses.


Ibid., p. 25.


Ibid. pp. 495-6, emphasis added.

[1869] LR 4 CP 739, p. 741, emphasis added.

All this raises a pressing question: If the proposition expressed by a sentence like ‘The defendant’s negligence caused the plaintiff’s harm’ depends on the modal base and ordering
source, what is the right modal base and ordering source to use for the purposes of determining an agent’s (degree of) responsibility or liability for the harm? I explore this question more fully elsewhere (see my ‘Contextualism and the Normative Connections Problem’).

25 For a good introduction to this debate, see Stegmann (2012).

26 This is plausibly what Sterelny and Kitcher have in mind when they describe genes as sufficient for many biological phenomena “relative to any standard environment” (Sterelny and Kitcher 1988, p. 349; my emphasis).

27 According to Lewis, for example, the worker and the manager “disagree only about which part of the causal history is most salient for the purposes of some particular inquiry” (Lewis 1986b, p. 215).

28 Does this view imply that the causal facts would have been different if the moral facts had been different (and isn’t that absurd)? No, at least on a natural reading of this counterfactual. Compare with gradable adjectives again. Suppose I’m in a context where ‘Kim is tall’ expresses the false proposition that Kim is tall for a person. Now consider the sentence ‘Had everyone else been shorter, Kim would have been tall’. There’s a true de dicto reading of this counterfactual: In the relevant counterfactual worlds, Kim is tall compared to other people in those worlds. But the natural reading is the false de re reading: In the relevant worlds, Kim is tall compared to actual people (this latter reading is false, since Kim’s height wouldn’t have been any different if everyone else had been shorter). In this case, then, it’s natural to interpret the standard of comparison as outside the scope of the modal operator. Similarly for causal claims. Suppose we’re in a context where ‘The pay-rise caused the bankruptcy’ expresses the proposition that the pay-rise caused the bankruptcy morally speaking, and suppose that this proposition is false. Then the natural reading of ‘If the moral facts had been different, the pay-rise would have caused the bankruptcy’ is the false de re reading: In the relevant counterfactual worlds, the pay-rise caused the bankruptcy relative to the actual moral ordering source. This reading is false, since the counterfactual structure of the case wouldn’t have been any different if the moral facts had been different. Thanks to Jake McNulty and Caleb Perl here.
Necessary Connections in Context

Alex Kaiserman

29 ‘Roughly speaking’, because this simple counterfactual analysis runs into problems in familiar cases involving redundant causation. These problems won’t be relevant in what follows.

30 Hoffmann (2011, p. 4) goes as far as to claim that “no judge in fact adopts” Stapleton’s recommendations, despite their influence in philosophical circles. Similarly, having acknowledged that “[t]he causal parity thesis derives its initial plausibility from the contemplation of the enormous causal complexity of the processes that constitute life at the molecular level”, Weber notes that it is “somewhat ironic” that “we would know close to nothing about this complexity if it wasn’t for gene-centered research” (Weber forthcoming).

Acknowledgements: Thanks to John Hawthorne, Ofra Magidor, Delia Graff Fara, Laurie Paul, Boris Kment and an anonymous referee for comments on previous drafts of this paper. Thanks also to audiences in Austin, New York, Princeton and Stockholm for helpful questions.

References


