Note to Readers: This paper, minus the Appendix, will be forthcoming in *Advances in Experimental Philosophy of Causation* (ed. Willemsen and Wiegmann) Bloomsbury Press. The Appendix added to this version contains some material that will be omitted from the published version for reasons of space but which I thought might be of some interest.

**Mysteries of Actual Causation: It’s Complicated**

**James Woodward**

**HPS, University of Pittsburgh**

1. **Introduction**

This chapter explores some issues having to do with actual cause judgments—judgments that some particular event caused another, as in “the impact of the rock caused the shattering of the bottle” or “Smith’s smoking caused his lung cancer”. This topic has received a great deal of attention recently in a number of different disciplines. In philosophy, computer science and psychology there are a number of more or less formal models of such judgments, often (but by no means always) making use of directed graphs and structural equations. (e.g. Hitchcock, 2001, Woodward, 2003, Halpern & Hitchcock, 2015, Halpern, 2016, Icard, Kominsky & Knobe, 2017) These models can be understood either as *descriptive* (as attempting to represent and in some cases to explain the judgments people make) or, in some cases, as *normative* (as characterizing the actual cause judgments people *ought* to make). In a number of cases they have *both* descriptive and normative aspirations—for entirely defensible reasons, as I explain below. At the same time there is an ever-expanding body of empirical work, from psychology and experimental philosophy on the people’s actual cause judgments.

Actual cause judgments (although not always under that name[[1]](#footnote-1)) have also been a topic of ongoing interest in more traditional philosophical discussion. Here there tends to be less use of formal models but there are still many attempts to capture or represent actual cause judgments via a mixture of verbal and partially formal proposals. (For a recent example see, Paul & Hall, 2013.) In this literature, the judgments to be captured are often described as “intuitions” that the writer has about various scenarios. For example, a scenario (EX1.1) may be constructed in which Suzy and Billy throw rocks at a bottle, Suzy’s rock strikes the bottle first, the bottle shatters and Billy’s rock passes through the empty space previously occupied by the bottle. The writer then announces that she has the intuition that the impact of Suzy’s rock caused (where this is understood to mean “was an actual cause of ”) the shattering and Billy’s rock did not. Some philosophers think of intuitions of this sort as potential sources of information about the underlying metaphysics of causation or about “causal reality”, but whether or not they claim this, they at least try to develop accounts that reproduce such intuitive judgments, while perhaps satisfying other constraints which may be more metaphysical in nature. Often it is assumed if they can produce an account that captures a wide range of intuitive judgments (and perhaps additional constraints), this shows that the account in question is (in some sense) warranted or correct. Similar assumptions seem to be made in some treatments of actual causation developed by non-philosophers. For example, in Halpern’s (2016) book length treatment, the standard for correctness also seems to be largely agreement with intuitive judgment.

One obvious question raised by these various literatures (and accompanying research agendas) is how if at all, they might fit together. For example, how should we think about the relationship between normative models of actual cause judgment and descriptive facts about how people in fact judge? What is (or should be) the relationship between empirical investigations of the sort conducted by psychologists and experimental philosophers into people’s patterns of judgments and the “intuitions” about causal relationships to which some philosophers appeal?

The rest of this essay is organized as follows. Section 2 provides an overall framework for thinking about causation, including actual cause judgments. I argue that causal claims of all varieties should be understood in terms of the goals or functions that they serve. This framework has been fairly successful in characterizing so-called type causal claims but is less successful in dealing with actual causation, in part because the function or functions of such judgments remain controversial. More specifically, I argue that we have a somewhat plausible functional account of actual cause judgments that involve *causal selection* but this account does not seem to fit actual cause judgments that are apparently not influenced by selection considerations and instead involve a distinct notion that I call “actual causal contribution”. At least some of the failures of current theorizing derive from attempts to provide a single unified account of what may be rather different varieties of actual cause judgments with different functions. Section 3 argues, contrary to what some metaphysicians seem to assume, that actual causation (however this is understood) is not a foundation or ground for all other sorts of causal judgments. Instead judgments of actual causation have a number of distinctive, even idiosyncratic features not shared by other varieties of causal judgment. Section 4 discusses the relationship between causal selection and actual cause judgments. Section 5 describes a standard account of actual causation that draws on the resources of structural equations and directed graphs. Section 6 describes several counterexamples to this standard account and an alternative more recent treatment, due to Halpern (2016), that attempts to accommodate the counterexamples by incorporating “normality” considerations. I argue that Halpern’s account tracks our practices of causal selection fairly well but is less successful in characterizing non-selective judgments of the actual contribution variety. Sections 7-9 attempts to flesh out this actual contribution notion in more detail. Section 10 contains some methodological reflections concerning the distinctions described earlier and also regarding the use of verbal probes to explore people's actual cause judgments. Section 11 speculates on a possible function for actual contrbution judgments and Section 12 concludes.

1. **Framework and Background**

Actual cause judgment is, in my view, just one possible variety of causal judgment among many (in fact, as we shall see, a kind of causal judgment with some rather idiosyncratic features). Like other forms of causal judgment, actual causal judgments are best approached from a “functional” standpoint. (Woodward, 2014, 2021). What I mean by this is that causal judgments of any variety should be understood in terms of their function, goal or point —what people are trying to do or accomplish when they make such judgments. This focus on function in understanding causal cognition is now fairly widely, although by no means universally, accepted—see. for example, Icard et al., 2017. As an interventionist about causation (see Woodward, 2003), I take one of the most distinctive and important functions of causal thinking to have to do with discovery and use of information relevant to manipulation and control. But there are many different kinds of manipulation-related questions and different sorts of causal judgment can be thought of as answers to such different questions. For example, a so-called type-causal claim such as “smoking causes lung cancer” suggests that intervening on whether people smoke will change the incidence of lung cancer—type-causal claims can be thought of as having the goal or function of providing information relevant to questions of this sort. As we shall see, actual cause judgments can also be thought of as answers to rather different but nonetheless manipulation-related questions.

Although functional treatments of type-causal claims have been fairly successful, I think it fair to say that similar treatments of actual cause judgment have so far been less illuminating. This is at least in part because it is a disputed issue what the goal(s) or point of actual cause might be or whether this is even an appropriate question to ask about such judgments. It may also be the case, as I will suggest below, that actual cause judgments are a heterogenous class, with different judgments having different goals or functions. One candidate for such a function, advanced by a number of researchers (e. g., Hitchcock & Knobe, 2009, Icard et al., 2017), is that it has to do with the identification of the most effective targets of intervention in some specific actual situation, given that one wants to change some effect. That is, *c* is judged to be an actual cause of *e* when intervening on *c* is thought to be an effective way of changing *e* (or a more effective way of changing *e* than other possible intervention targets), given the actual situation in which *c* and *e* obtain. I will explore this suggestion below, suggesting that it fits some sorts of actual cause judgments better than others

I distinguished above between normative and descriptive treatments of causal judgment. One obvious normative standard (and the one assumed in this essay) comes from the notion of function just described: one shows that some account of causal reasoning, judgment, learning etc. is normatively good (correct etc.) by showing that proceeding in accord with the account contributes to or promotes (is an efficient means to) various goals associated with causal thinking. Applied to actual cause judgment this means that if we are to develop an account that tells us which such judgments are “correct” or “warranted” we should proceed by identifying goals or functions associated with such judgments and then use these to distinguish judgments that contribute to these goals from those that do not.

Of course, even given this understanding of “ought”, how people ought to judge and how in fact they judge are two different matters. But the normative and descriptive can be brought into closer contact if an additional assumption holds: that people are, at least a significant portion of the time, “rational” or behave in a normatively appropriate way in the causal judgments they make. (“Rational” here should be understood in purely means/ends terms—some pattern of judgments is “rational” to the extent that it contributes to rather than frustrating ends that the judgers have, where these have to with e.g. identifying relations relevant to manipulation and control, among other possibilities.) To the extent that this is so we can use normative models to capture and explain features of judgments that, as an empirical matter, people make. At the same time if we find that people’s causal judgments exhibit, as an empirical matter, certain features and not others this can suggest (*suggest*, not by itself conclusively establish) various hypotheses about their ends or goals, and the ways in which their judgments are effective means to these.

As an illustration, one aspect of actual cause judgment (although arguably not all that is involved in such judgments—see below) concerns *causal selection*. Suppose that two events—*c1* and *c2*–must both be present for *e* to occur (the causes act “conjunctively) and that together they are in some relevant sense causally sufficient for *e*, as when striking a match (*c1*) in the presence of oxygen (*c2*) is followed by ignition *e*. In such cases, when subjects are asked what caused *e* they tend to “select the “abnormal” event *c1* rather than the more normal event *c2*, and when they are asked to make “causal strength” judgments they assign higher ratings to *c1* than to *c2*. A number of different verbal probes have been used to elicit such judgments. Knobe and Fraser (2008) and Icard et al., (2017) ask subjects to rate on a 7-point scale whether they agree with descriptions of the form “*c* caused *e”.* In the former study, assignment of a higher rating to *c*1 than to *c*2 is taken to indicate “selection” of *c*1. In the latter study, stronger agreement is taken correspond to a judgment of greater causal strength. Vasilyeva et al. 2018 use a somewhat different probe to measure causal strength: “ how appropriate is it to describe this relation as causal? ” As noted below, variations in wording can make a difference to the judgments elicited in ways that are not fully understood[[2]](#footnote-2).

Assignment of higher causal strength to abnormal causes in the conjunctive case makes sense if (as suggested earlier) one of the functions of such causal selection judgments (or one of the goals subjects have in making such judgments) is to identify the events or factors that are effective targets of intervention in the actually obtaining situation. As a standard example, described but not tested in Icard (2017) oxygen is almost always present on the earth’s surface (as reflected in the judgment that its presence is normal) and in most cases it is difficult or impossible for people to change whether it is present. By contrast, match strikings are comparatively unusual statistically speaking (abnormal) and relatively easy to control. If one wishes to bring about or prevent a match ignition, in most circumstances it makes sense to focus on whether the match is struck or not, rather than whether oxygen is present. This is not to say that the effectiveness of intervention rationale applies to all cases of actual cause judgment, but it does illustrate one possible version of functional approach to actual causal judgment.

Next some remarks about the role of appeals to “intuition” which, as I have said, is ubiquitous in discussions of actual causation. As argued elsewhere (Woodward, 2021) I think it is implausible to view intuitive judgment as a kind of rationalistic grasping of underlying metaphysical truths or anything similar. Instead, to the extent that there is genuine information in reports of intuition these should be understood as claims about how the intuiter and others judge—in other words, such reports are simply empirical claims (or evidence for empirical claims) about practices of causal judgment. Consider the episode (EX1.1). When we have the “intuition” that the impact of Suzy’s rock caused the bottle to shatter, this simply reflects our judgment about the case with (in most cases, if the intuition is of any interest) the further claim that others will judge similarly. It is of course an empirical question whether this further claim is true, although it is very plausible that it is[[3]](#footnote-3). If we think of the significance of intuitive reports in this way, the information they potentially provide is broadly similar to the information provided by experimental investigations into patterns of causal judgment conducted by psychologists and experimental philosophers. Insofar as philosopher’s intuitions provide information of this sort, there is nothing special or unique about them.

Traditional “armchair” philosophers often dismiss the relevance of empirical studies of folk judgment to their projects, presumably in part because they expect (or fear) that the empirical studies will produce results in conflict with their intuitions. However, my assessment is that this expectation of conflict has largely turned out to be incorrect when the topic is causal judgment. That is, as an empirical matter, philosopher’s intuitions (when construed as claims about how people will judge) often turn out to produce results that are consistent with more systematic empirical studies[[4]](#footnote-4)—something that ought to be a source of satisfaction to the armchair oriented. This is not to say that the empirical studies are superfluous. For one thing they can serve as additional checks on philosopher’s reports of intuitions. More interestingly, there are many issues concerning causal cognition that can be explored by empirical methods employed by psychologists but not by armchair appeals to intuition[[5]](#footnote-5). The empirical methods can also yield information that might possibly be provided by armchair methods but which traditional philosophers have not tried to provide. For example, as remarked above, it is standard in psychological research to ask subjects to make graded judgments of “causal strength” rather than binary judgments (did *c* cause *e*?). Philosopher’s appeals to intuition tend to involve only the latter. Finally, there are patterns of judgment discovered in psychological experiments that are at least somewhat surprising and that have not been uncovered by armchair methods—one example is “abnormal deflation” in disjunctive causal systems[[6]](#footnote-6).

Given this continuity (in many cases) between philosopher’s intuitions and experimental results, I will draw on both in what follows. To the extent that philosophical intuition is empirically accurate as a characterization of people’s judgments, it is obvious why it has a role to play in descriptive accounts of such judgments. However, following the program outlined above, intuitive judgments can also play a legitimate role in normative inquiry. To the extent that we can assume people reason and judge in normatively appropriate ways, empirical results about how they do reason can *suggest* hypotheses about how one ought to reason and judge and functions and goals of such reasoning. I write “suggest” because on this view, agreement with intuitive judgment is not by itself *sufficient* to establish normative correctness. Instead establishing normative correctness requires the kind of means/ends analysis and reference to function described earlier. For example, if it is correct or warranted to select the occurrence of the short circuit rather than the presence of oxygen as “the cause” of the fire, this is so not because most people judge in accord with this selection but rather because, assuming the effectiveness of intervention goal or function, intervening on the short circuit better conduces to this goal. In fact, assuming the effectiveness of intervention rationale, a number of actual cause judgments that people regularly make seem, if not mistaken, highly suboptimal. For example, in making actual cause judgments about causes of disease and industrial accidents, most people focus on “proximate” causes, such as infection or operator error, even though the most effective targets of intervention are often more distal, having to do with social organization.

There are many other reasons why assessing theories of causal judgment merely by fit with intuitive judgments (without attention to function) is unlikely to produce fully satisfactory results, either from a normative or descriptive perspective. In connection with actual cause judgment, I will mention just two. As Glymour et al. 2010 note, the examples of actual cause judgments explored by philosophers represent a tiny subset of the possible cases. Even for these known cases it is far from clear that any of the extant formal models fully capture our intuitive judgments. (Again, Glymour et al, 2010) Moreover, even if we were to find a model that captures our judgments in all of the examples considered so far, this by itself provides no assurance that will not be other examples, perhaps many, for which the model fails. Adding functional constraints, can help with this problem: If we specify a function for actual cause judgments and a formal model characterizing how actual cause judgments meeting the conditions of model achieve this function *and* if we find that the model captures the judgments people actually make (or some substantial subset of these) this may help to assure us that there are not a large number of unexplored counterexamples lurking around the corner or, if there are, that these need to be understood in terms of some other function.

1. **The Place of Actual Causation**

Iturn next to some general remarks about the place of actual cause judgment in the overall panoply of practices of human causal judgment. My general picture is that there are structures or patterns of counterfactual-supporting dependency relations that are “out there” in the world. These are present at different levels of organization, from the behavior of elementary particles to cells to people and so on. Various sciences—physics, chemistry, biology, economics—provide information about these structures and this information allows us to answer questions about what would happen if various interventions were to be performed. Causal models and structural equations of the sort described below encode portions of this dependency information.

Different kinds of causal judgment attempt to report on different features of this underlying structure—they attempt to provide answers to different kinds of queries we may make concerning this structure. What is distinctive about actual cause judgments is the kind of questions they attempt to answer and the information they draw on in doing so.

To spell this out, let us return to the often-invoked contrast between actual cause and so-called type level claims. Although I have employed this terminology, it is potentially misleading. It is not plausible to interpret (3.1) “Smoking causes lung cancer” as asserting that some uninstantiated abstract type or property (“smoking”) stands in a causal relation to some other abstract property (“lung cancer”). Instead (3.1) is best understood as the claim that when particular individuals smoke this can or will cause those individuals to develop lung cancer. Both in the case of (3.1) and (3.2) “Jones’ smoking caused his lung cancer”, any causation that goes on is at the token level—in both cases the units, so to speak, are individual episodes of smoking and lung cancer, co-attributed to particular people. (3.2 ) differs from (3.1) in that it purports to answer a different question about this underlying structure—among other things, (3.2) picks out or refers to a particular (extended) episode of smoking causing lung cancer while (3.1) does not.

Another important difference between (3.1) and (3.2), noted by a number of writers, is this: In assessing a claim like “Smoking causes lung cancer”, one begins with a putative cause—smoking—and reasons “forward” from this to determine whether it has some effect (e.g., lung cancer). The reasoning is thus from causes to possible effects. Although it would of course be highly unethical to conduct an experiment in which subjects are randomly assigned to either a treatment group in which they are forced to smoke or, alternatively, to a control group in which they are prevented doing so, the results of such an experiment would provide us with very good evidence for whether claim (3.1) is true. Similarly, we might establish that (3.1) is true by applying some appropriate causal modeling technique to statistical information about patterns of covariation involving the incidence of smoking, lung cancer and other variables in a population

Actual cause judgments are associated with a very different question and present us with a different inferential task, requiring different evidence. Typically, when one makes such judgments one is presented with a known effect such as the occurrence of lung cancer in a particular individual, Jones. It is assumed that this effect has some cause or causes. The task is to reason “backwards” from this effect to what caused it—was it Jones’ smoking, or something else such as his exposure to asbestos or radon? Or was this perhaps a case of overdetermination involving two independently acting causes each of which was sufficient for the effect? In this case, the direction of reasoning is the opposite from what it is in connection with (3.1). Although experimental evidence can certainly be relevant in various ways to (3.2) one can’t settle whether (3.2) is true by doing an experiment, however well-designed. Similarly, population level statistics and causal modeling will not by themselves provide information sufficient to determine whether (3.2) is true.

In cases in which several different candidates for the cause of the effect of interest are present (Jones both smokes and was exposed to asbestos), determining which of these was the actual cause typically requires further specific information about the system of interest (in addition to generic causal information relating smoking to lung cancer and asbestos exposure to lung cancer). In a number of cases this takes the form of information about the presence of a “mechanism of action”. The idea here is that when a cause is not merely present but efficacious in causing an effect (that is, an actual cause of the effect), there may be independent evidence (that is, evidence in addition to the occurrence of the cause and the effect) that this is the case or, alternatively, evidence that despite being present the cause did not produce the effect. This can take the form of evidence about mediating variables (a “connecting process”) that must be present if the cause is to produce the effect, evidence about the presence of other effects (but not mediating variables) that will be present if the cause is efficacious and information about the spatio-temporal relations between the candidate cause and effect. One sometimes speaks in such cases of finding a “signature” or “fingerprint” that one candidate actual cause and not another was operative in producing an effect. In rocks thrown at bottles scenarios, unless something very unusual is going on, even if a number of rocks are “present”, the rock that is the actual cause of the shattering is the one that comes into spatial contact with the intact bottle. If Jones is shot in the head and dumped in the river, forensic analysis may be able to determine (e.g., based on the presence or not of water in his lungs) whether the actual cause of his death was drowning or the gunshot.

My claim is *not* that that such information about mechanisms of action can be used to define or characterize what it is for one event to be an actual cause of another—I see little promise in such a project. Rather, my point is that actual cause judgments are answers to different questions than those to which other sorts of causal claims are directed and also rely on different sorts of information. As noted below, this has implications for whether information contained in actual cause claims will fix the truth values for other sorts of causal claims (or conversely).

In this connection it is also worth remarking that some of the standard examples used to illustrate actual cause claims in the philosophical literature suggest a misleading picture about how easy it is to establish such claims. When Suzy’s rock comes in contact with the bottle before Billy’s it is straightforward to use this information to establish the actual cause of the shattering. But many actual cause judgments in which we are interested are not like this. Establishing that global warming was an actual cause of some particular weather event, like a hurricane although arguably not impossible, is highly non-trivial. Similarly, it can be very difficult to establish that it was exposure to asbestos rather than smoking that caused Jones’ lung cancer – something exploited by lawyers for industries using asbestos in tort actions.

A great deal of theorizing in science does not aim at the explanation of particular events via actual cause claims but rather at accounting for more generic explananda—patterns or phenomena in the sense of Bogen and Woodward (1988). For example, General Relativity provides explanations of such phenomena as the deflection of light by massive bodies, the Hodgkin Huxley model explains the generic shape of the action potential and so on. In such cases it may be assumed that all particular episodes of light deflection etc. have the same explanation so that in discovering the correct explanation we are again faced with a rather different inferential problem than that of choosing among different candidates for the actual cause of some particular outcome. Even when different possible causes of an outcome may be operative, in scientific investigation the interest is often in whether there is a pattern of outcomes that suggest a generic causal relation, rather than identifying the actual causes of particular outcomes. For example, in a randomized controlled trial that tests the efficacy of a drug in causing recovery from an illness, what is of interest is usually whether there is a higher incidence of recovery in the treatment group than in the control group. If there is, this implies that some individuals in the treatment group were caused to recover by the drug, but the experiment as described does not identify these individuals and it may be difficult or impossible to do so. Moreover, there may be little point in doing so if what is of interest is a generic claim concerning whether the drug causes recovery.

This is not to say that scientific investigation is never concerned with establishing actual cause claims. Scientific investigations into particular historical episodes such as the Cretaceous-Tertiary extinction of the dinosaurs are naturally viewed as concerned with establishing actual cause claims. Moreover, actual cause claims are of central interest in many other contexts and disciplines—not just in common sense contexts but in law, in diagnosing the causes of failure in engineering contexts (what caused the challenger disaster?) and in history more generally. My point is simply that actual cause claims play a much less central role in science than many suppose.

I stress this because there is a tendency among some philosophers—particularly those with a metaphysical orientation—to think that actual cause claims, largely as revealed in our ordinary practices of intuitive judgment, provide us with a notion of causation that is a foundation for *all* causal judgments[[7]](#footnote-7). This may be thought to follow from a metaphysical thesis according to which other sorts of causal claims are “grounded” in actual causation. I will not try to engage with this metaphysical claim but the considerations described above (as well as others discussed below) suggest that it is a mistake to think that all of the features present in actual cause judgments will also be present in other forms of causal judgment. In other words, the common philosophical practice of undertaking to discuss “causation” and then discussing only actual cause claims as though these will have all of the features possessed by other kinds of causal claims is misguided.

1. **Causal Selection**.

One of many issues about actual causation raised by both the philosophical and psychological literature concerns the role of causal selection in such judgments. Philosophers with an interest in actual causation have had very different attitudes toward causal selection. Hall (2004), after noting the existence of such practices, says that his target is what he calls an “egalitarian” notion of (actual) causation. According to this notion, in the match ignition example, both the striking and the presence of oxygen are among the actual causes of the ignition and an account of actual causation should capture this fact. This egalitarian notion is seen as capturing an objective causal structure while selection of particular factors from this structure is seen as reflecting “subjective” and likely unsystematic factors having to do with our interests. Causal selection is thus a potential source of error or distortion in dealing with actual causation. This of course assumes that we possess (or can isolate) a notion of actual causation in which selection practices play no role and that we can recognize when our judgments track such a notion. I suggest below that it is an open question whether this assumption is correct.

More recently, there has been a great deal of interest in the details of how causal selection influences actual cause judgments, both among philosophers and psychologists. Indeed, to a considerable extent, current theories of “actual causation” now aspire to capture causal selection practices, often within formal frameworks. In other words, “actual causation” and “causal selection” are not distinguished in the way that Hall advocates but are rather seen as a single topic which should be given a unified treatment.

If this is the most appropriate way of approaching the topic of actual causation important consequences follow. As noted earlier a number of empirical investigations have shown that judgments of “normality” vs “abnormality” play an important role in causal selection. In “conjunctive” structures like the match example, the statistically abnormal cause (the striking) tends to be selected as the cause and also receives a higher causal strength rating than the more normal cause Moreover, this pattern extends to causes that involve norm violation vs norm conformity. In a scenario described in Knobe and Fraser (2008), staff are permitted to take pens from the secretary’s desk while faculty are prohibited. Both a staff and faculty member take pens from the desk with the result that the secretary is unable to take an important message. Here the norm-violating action of the faculty member receives a higher causal strength rating than the action of the staff member.

Presumably, no one thinks that causal selection judgments that reflect norm violations are tracking some fundamental objective structure in the world. At least in many cases generic causal claims in the physical sciences (such as those mentioned above) do not appear exhibit this sort of norm sensitivity. Somewhat more subtly, according to many normative theories, whether a generic causal relation like “*C* causes *E*” holds and how “strong” it is or whether it is appropriately described as causal should not depend on how frequently *C* occurs[[8]](#footnote-8). Moreover, as a descriptive matter, at least in many cases, when we are dealing with generic causal relations judgments of causal strength are *not* influenced by the frequency with which *C* occurs or at least not influenced in the way in which actual cause judgments are. Here again the factors that influence such generic causal judgments appear to be rather different from those that influence causal selection in actual cause judgments.

If causal selection judgments are viewed as a legitimate aspect of actual cause judgments or if it is not possible to separate out the selection aspect from a more egalitarian core to such judgments in the manner Hall suggests, this strengthens the case that such judgments are not metaphysically fundamental (or fundamental to understanding causation in science) in the way some have supposed, given the differences between selection judgments and other, more generic causal judgments just described. However, the issues here are complicated. Returning to some of the actual cause judgment discussed above, it is not clear all of them involve the kind of causal selection effects that are present in the short circuit/oxygen or pen example. For what it is worth, my own intuitive judgment (and what I assume will be the similar judgment of others) in EX1.1 that the impact of Suzy’s rock caused the shattering and Billy’s did not does not seem to be sensitive to normality considerations in the way in which causal selection judgments are. If Suzy’s throw was very normal and Billy’s very abnormal (Suzy frequently throws rocks at bottles and Billy does not, Suzy is permitted to do this and Billy is not) or vice-versa, this does not make me any more or less willing to judge that Billy’s throw was the actual cause if Suzy’s rock came into physical contact with the bottle and Billy’s did not[[9]](#footnote-9). (I acknowledge though that my judgment may be eccentric—more systematic investigation is needed.) A similar assessment may be warranted for a number of the other examples described above, such as those involving disease causation—e.g., the normality or not of Jones’ asbestos exposure does not seem to influence our judgment about whether this was an actual cause of his lung cancer.

Considerations of this sort may seem to suggest that there is, after all, a notion of actual causation and actual cause judgment that is distinct from and independent of whatever is involved in causal selection judgments and thus that it is a mistake to try to construct a single unified theory of both kinds of judgment. Instead we have two kinds of judgments, one reflecting something more like Hall’s egalitarian notion and the other reflecting selection effects. I tentatively endorse this claim below, describing the non-selective notion as involving “actual causal contribution”. On the other hand, as we shall see, several theorists, proceeding on the assumption that we should look for a single unified theory of actual causation, argue instead that various examples force formal theories of actual causation to incorporate normality considerations (and hence what appear to be selection effects) into our understanding of *all* actual cause judgments. These theorists implicitly reject the idea that there are two (or more) different varieties of actual cause judgment. Moreover, there are cases in which it is by no means clear whether we are dealing with causal selection judgments or some non-selective more egalitarian notion. So the idea that there is a sharp separation between actual cause judgments that involve normality influenced causal selection and those that do not, although intuitive, may not prove to be sustainable.

1. **Structural Equations and Formal Models of Actual Causation**

To explore these issues in more detail, I turn to more detailed remarks about some formal theories of actual causation and the confrontation of these with experimental results and philosopher’s intuitions. One currently popular form such models take makes use of structural equations and directed graphs (e.g. Halpern, 2016). For reasons of space and because I assume that such models are familiar to many readers, my exposition of these will be brief and unrigorous. In structural models, generic causal relations are represented by equations relating variables (or more pedantically, whatever in the world corresponds to variables) with, by convention, variables representing effects occurring on the left-hand side of such equations and variables representing the direct causes of those effects occurring on the right- hand side. A natural way of understanding such equations is that they encode information about what would happen if various interventions were to be performed. Thus

(5.1) Y= 2X1 + 3X2

is interpreted as claiming that an intervention that changes the value of X1 by dx1 (while leaving the other right-hand side variable X2 unchanged) will be associated with a change in the value of Y of 2dx1. The most straightforward way of representing actual cause claims within such a framework (and the one that I will adopt for purposes of this essay) is to take them to relate *values* of variables, where the variables themselves stand in relationships specified by the structural equations. For example, in connection with a structure like (5.1) one might ask whether, in some specific situation, X1 taking the value = 1 is an actual cause of Y= 2. (I will say more shortly about how this might work.) We can also associate a *directed graph* with such systems of structural equations, where the rule is that each equation corresponds to a structure in which there is a directed edge from each of the right-side variables into the left-hand side variable.

The use of the structural equations in the form described commits us to the claim that the generic causal relations modeled are deterministic. It is also generally assumed that the systems modeled do not contain causal cycles. I will adopt both assumptions in what follows. It is also worth noting that structural equations and graphical models were originally developed to represent generic causal relations; their extension to actual causation is a relatively recent development, apparently beginning with Halpern and Pearl (2005). The extent to which these frameworks are suitable for this latter purpose ought to be regarded as an open question.

In most early versions of structural equation treatments of actual causation, the issue was approached in the following way: Suppose we are given (5.2) the full or complete set of structural equations governing the system of interest (that is, we proceed on the assumption that these are correct, so that these give us the ground truth about the relevant generic causal relations) and we are given (5.3) the actual values taken by all variables in the system of interest. The task was then taken to be this: to specify an algorithm or set of rules that takes the information in (5.2) and (5.3) and outputs (5.4) the actual cause relations that hold in the system. As noted above, the test for whether the output (5.4) is correct was commonly taken to be agreement with the intuitive judgments of the researcher or perhaps others, although more function-based standards of assessment are also possible.

The idea that there should be such an algorithm might seem to have considerable antecedent plausibility. First, in a considerable range of cases (although arguably not all, as we shall see) people do seem to have firm judgments about actual cause relations, suggesting some stable target to be captured. Moreover, if we think of (5.2) as specifying something like the “laws” or dependency relations governing the system of interest, it may seem very plausible that these, in connection with (5.3) the actual values taken by the variables characterizing in the system, should be sufficient to “fix” all of the actual cause relations in the system. That is, the thought is that (5.2) and (5.3) together specify all of the information relevant to the actual cause relations in the system—all of the information that goes into people’s intuitions—so that there must be some mapping from (5.2) and (5.3) to people’s intuitive judgments.

Philosophers (e.g. Lewis, 1986, Paul and Hall, 2013) who use so-called “neuron diagrams” to investigate actual cause judgment employ a somewhat different representational framework but (I think it fair to say) have broadly similar expectations. These diagrams represent connections or dependency relations between nodes and also the values of the nodes (whether they are “on” or “off”) —information that parallels the information about dependency relations between variables and their values in the structural equations framework. As in the structural equations representation, the goal of the neuron diagrammers is to specify rules that allow one to go from this information to the intuitive actual cause judgments one makes about such systems.

In what follows I will focus on the structural equations framework both because I am more familiar with it and because it is more widely used in current discussion. My guess is that much of what I say about this framework will transfer to (have analogues in) neuron diagram treatments, but I will not try to show this.

Having set the common problem up in this way, I can now come to an important punchline: it is, at the very least, extremely difficult to fully carry out the program just described. Although there are a number of proposals relying on the information described above (dependency relations plus actual values) that successfully capture a number of our intuitive actual cause judgments, none seem to capture the full range of our judgments. Of course, one possible reaction is that if we just try harder and explore more possibilities, we will find an account that works. Another possible reaction—which I’m inclined to think must be closer to the truth—is that actual cause judgments (or at least some of them—see below) are influenced by other inputs besides those just described and that is the reason why the proposals fail[[10]](#footnote-10). In fact, as we shall see, several more recent modeling proposals in a broadly structural equations framework seem to acknowledge this.

To develop this theme, I turn to a very partial exploration of several modeling proposals beginning with one in the spirit of Halpern and Pearl. (See also Hitchcock, 2001 for a similar proposal.) As before my goal is not precision but conveying an underlying intuition. First some terminology. Given a directed graph G with variables V representing the system of interest, with X1 and Y in V, P is a *directed path* from X1 to Y if there is a set of additional variables X2, X3,..,Xn also in V such that X1🡪X2, X2🡪 X3,…Xn🡪Y. An *off -path variable* with respect to path P is any variable in G that is not on P. Then the proposal is

(**AC**) X1=x is an actual cause of Y=y if the following conditions hold:

**AC1**) X1= x and Y=y—i. e., these events actually occur.

**AC2)** There is a directed path P from X 1 to Y in G such that for some allowable setting or “contingency” (see below) of values of variables thatare off P, an intervention that changes the value of X=x1 to some alternative value X1= x1\* would change the value of Y=y1 to some alternative value Y=y\*1.

The specification of what counts as an allowable setting is a matter of some delicacy, with a number of different proposals in the literature. However, the intuitive idea or at least part of it is that the allowable off-path settings (contingencies) in **AC2** are those that do not change the value of Y or any other variable on the path P.

To illustrate how this is intended to work, consider several examples:

**Simple Pre-emption: (Ex 5.1)** If Suzy throws, her rock will shatter the bottle, Billy will not throw if Suzy does, but if Suzy does not throw, Billy will throw and his rock will shatter the bottle. Suppose that Suzy throws. With C1 representing whether or not Suzy throws, C2 representing whether or not Billy throws and E representing whether or not the bottle shatters, we have the following equations and associated graph.

E= C1 v C2

C2= not C1

Actual values: C1=1, C2= 0, E=1

C1

E

C2

C1🡪 E is a path P from C1 to E. C2 is an off-path variable with respect to P. Set C2 to its actual value = 0—setting a variable to its actual value is always allowable because this does not change the values of variables on P. Under this setting an intervention that changes the value of C1 from 1 to 0 would change E from 1 to 0. Hence, in accord with intuitive judgment, C1=1 is an actual cause of E=1. C2=0 is not an actual cause of E=1 because under the setting C1=1, changes in the value of C2 make no difference to the value of E and, the alternative possible setting of C1 (C1=0) is not permissible, since that would change the value of E.

**Over determination: (EX5.2):** Two riflemen simultaneously shoot (S1, S2) and hit victim. Either hit is sufficient for death (D)

Equation: D= S1 v S2

Actual values S1=1, S2=1, D=1.

Graph:

S1

D

S2

Consider whether S1 is an actual cause of D. AC 1 is satisfied. AC2 is also satisfied: There is a path P from S1 to D. S2 is an off-path variable with respect to P. Setting S2=0 is an allowable setting since setting S2 to that value will not change the value of D (D is still = 1). Under the setting S2=0, intervening to change the value of S1 will change the value of E, so S1= 1 is an actual cause of D=1. Parallel reasoning shows that S2 is also an actual cause of D=1. Although there is some dispute, this seems to be in accord with the intuitive judgments of most people; both shooters are actual causes of death.

Note that although AC succeeds on the examples above it plainly fails to capture judgments that reflect what we naively think of as causal selection. In the match/ oxygen example, the structural equation is:

F= M. O

with actual values, F, M, O= 1.

O is an off-path variable with respect to the path from M to F. Setting O at its actual value=1, the value of M makes a difference to whether F=1 so M=1 is an actual cause of F=1. But similar reasoning shows that O=1 is an actual cause of F=1, so that there is no asymmetry between the roles of S and O. This is not surprising if causal selection reflects normality judgments, since normality plays no role in **AC.** **AC** captures a notion of actual causation that looks more like Hall’s egalitarian notion.

Why does **AC** seem to work (to the extent that it does)? Here is one way of understanding the underlying rationale. (I introduce it now because I will make use of it later.) Suppose that causes (of any variety) are difference-makers, under some appropriate understanding of that notion -- that is, when *c* causes *e*, we expect that there is some change in *c* that in some circumstances will be followed by a change in *e* so that in this sense *e* depends on *c*. However, the difference-making aspect of a cause can be “hidden” when it is embedded in a larger causal structure, which is what happens in Ex5.1 and EX5.2. In EX5.1 the difference making role of Suzy’s throw is masked by the fact that Billy will shatter the bottle if she does not. We can uncover Suzy’s role if we “control for” or remove the elements that hide it. Obviously, though, we don’t want to change the values of variables lying on the candidate path by which *c* has a difference-making impact on *e*, since this will alter or disrupt that path. So we consider whether *c* has a difference-making impact on *e*, given settings of variables that are off the path under consideration. However, there needs to a limit on which off-path contingencies are allowed. In particular, when we consider whether there will be changes in *e* under changes in *c* for possible settings of the off-path variables we don’t want those change in *e* to be due to the settings of the off-path variables rather than the changes in *c*. The requirement in **AC2** that the off-path settings do not change the value of *e* or other variables on P from what they are in the actual situation attempts to accomplish this.

Let me note for future reference that the rationale just described is most compelling when a kind of invariance or robustness assumption holds regarding the *c🡪e* relation. That is, it is when we can assume that *c*’s difference-making role with respect to *e* can hold somewhat uniformly across a range of different situations (corresponding to settings of off-path variables), that the strategy of making that role appear by setting those variables to different values appears most plausible. Such an invariance assumption seems to be satisfied in EX 5.1 and EX5.2: Rock impacts have a shattering effect on bottles across a range of difference circumstances and we think that they can continue to have such an effect even when back-up causes like Billy’s throw are present, although in this case the difference-making role of the impact of Suzy’s rock is hidden. In other words, the causal role that the impact that Suzy’s rock would have if Billy were absent remains the same when Billy is present. In EX5.2, if each shot is sufficient for death, and there is no interaction between the shots, it is extremely natural to think of the causal contribution of S1to D as the same whether or not S2 is also present, which motivates bringing out the difference-making role of S1 by considering a situation in which S2 is absent. Thus the application of **AC** to the example again relies on some such invariance/robustness assumption. At least some of the limitations of **AC** that are explored below seem to involve cases in which such robustness assumptions don’t seem to hold.

1. **Problems with AC: Normality, Elections and Other Scenarios**

Although proposals along the lines of **AC** successfully capture a range of intuitive judgments, more recent discussion has uncovered a number of examples in which this is apparently not the case. Among other possible deficiencies, it appears that **AC** is overly permissive in the sense that it counts events as actual causes when they are not. (I write “appears” because whether this assessment is correct depends in part on issues having to do with the place of causal selection in a theory of actual causation. It would perhaps be more accurate to say that **AC** counts events as actual causes that would not readily be selected in a causal selection task.) I will describe two illustrations, due to Livengood (2013) and discussed in Halpern (2018).

Suppose Jack and Jill are potential voters in an area that votes heavily Republican. Jack leans strongly Republican but does not vote because he thinks that the Republican candidate will win without his vote. Jill would vote Democratic if she were to vote but decides not to vote because she is disgusted with the process. As expected, the Republican candidate wins overwhelmingly. With V1…Vn, representing the decisions of eligible voters (which can take values = R, D or abstain) and O the outcome, the graph consists of edges from each of V1,.. Vn into O. (We thus assume that the voters don’t influence one another’s votes.)

According to **AC** the abstentions of both Jack and Jill are actual causes of the outcome. This is because among the permissible settings of off-path variables are those that switch a number of the Republican votes to Democratic, as long as this is consistent with the Republican winning. Suppose there are 2n+1 eligible voters, 2n-1 of whom vote either R or D, with only Jack and Jill abstaining. Switch votes among those who vote so that there are n R votes and n-1 D votes. Then if Jack had voted Democratic there would have been a tie and the Republican would not have won. On the other hand, if Jack had voted R the Republican would have won. Under this contingency, Jack’s choice is difference-making. So Jack’s choice = abstention is an actual cause of the Republican winning. A parallel argument shows that Jill’s abstention is an actual cause of the Republican winning. Halpern claims that our intuitive judgment, though, is that although Jill’s abstention is such an actual cause, Jack’s is not. It is not clear to me that this is correct but it does seem wrong to treat Jack and Jill’s abstentions in exactly the same way (as **AC** requires), given the different ways they would have voted had they voted.

A second example, a slight variant on one also suggested by Livengood, is described by Halpern (2018) as follows:

consider a vote where everyone can either vote for one of three candidates. Suppose that the actual vote is 17–2–0 (i.e., 17 vote for candidate A, 2 for candidate B, and none for candidate C). Then not only is every vote for candidate A a cause of A winning, every vote for B is also a cause of A winning. To see this, consider a contingency where 8 of the voters for A switch to C. Then if one of the voters for B votes for C, the result is a tie; if that voter switches back to B, then A wins (even if some subset of the voters who switch from A to C switch back to A).

As Halpern observes, this may seem particularly unreasonable because if option C (which received no votes) had not been present, then votes for B would not be causes of A’s winning.

What these cases may seem to suggest is that at least some actual cause judgments are influenced by assessments of how likely (or perhaps “normal”) various alternatives to the actual situation are and that this in turn influences which settings of off-path variables are permissible, assuming that we continue to think in terms of a proposal that is broadly like **AC**. In particular, it is likely that if Jill had voted, she would have voted Democratic, so a scenario in which that happened seems or may seem relevant to our causal assessments. By contrast, if Jack had voted, it is extremely unlikely he would have voted Democratic, so considering a contingency in which that happens seems less appropriate than a contingency in which Jill votes Democratic. In addition, if the area is heavily Republican, a possibility in which a large number of voters switch from Republican to Democratic seems very unlikely. In the second example, given that C received 0 votes, an alternative in C receives 8 is naturally viewed as very unlikely. Again, one might think it follows that this an inappropriate or relatively inappropriate contingency for the purposes of applying **AC**.

Halpern’s proposal for capturing the role of these considerations is to add a “normality ordering” to a model along the lines of **AC**. This is a partial order which ranks alternatives to the actual situation in accord with how “normal” they are—e.g., Jill’s voting Democratic is more normal than Jack’s so voting. In Halpern’s proposal the normality ordering incorporates information about the probability of various variable values but it also can incorporate information about other sorts of norms, such as prescriptive norms or institutional rules. In this case, contingencies including actions that are more in accord with prevailing norms are viewed as more normal than those that violate such norms. This allows him to capture cases like Knobe’s pen example.

Given this ordering, normality considerations can be incorporated into causal judgment in at least two ways. One possibility is to add to **AC** the following restriction on permissible settings of off-path variables: alternative settings which change actual values to more abnormal values are not permitted while changes in the direction of greater normality are, with actual causation being assessed accordingly. (This might be expanded to prohibit other sorts normal to abnormal changes—for example, in the governing equations themselves – but I will not consider such possibilities here.) This proposal leads to categorical judgments according to which some *c* either is or is not judged an actual cause.

Another possibility is this: As noted above, in research on the empirical psychology of causal cognition, it is very common to ask subjects to make graded judgments of causal strength. Thus a theory of actual causation might itself take a graded form and be assessed in terms of how well it captures such strength judgments. That is, the theory would say that *c* is an actual cause of *e* to a greater or lesser extent (or an actual cause of greater or lesser strength), depending on how abnormal the contingencies have to be to reveal the dependence of *e* on *c*. There is an obvious sense in which this second, graded alternative seems more natural, since normality itself is a graded notion. Many recent discussions, including Halpern’s (2016, 2018) as well as Hitchcock and Knobe (2009) have taken such a graded notion of actual cause judgment to be the target to be captured. Adopting this graded notion, both Jack and Jill are relatively weak (non-strong) causes of the Republican victory since a contingency in which their votes make a difference to the victory would be “abnormal”[[11]](#footnote-11). Nonetheless, Jack’s abstention is an even weaker (albeit only slightly) cause of the Republican victory than Jill’s abstention since a contingency in which Jack votes Democratic is more abnormal than Jill’s so voting. (The proposal thus captures the intuition that Jill’s role is in some way different from Jack’s.)

To the best of my knowledge there have been no empirical studies of ordinary subject’s actual cause judgments in the specific election scenarios described above[[12]](#footnote-12). (Researchers have instead relied on their own intuitive judgments). For what it is worth, my guess is that to the extent that we have in mind a dichotomous notion of actual causation, according to which the only alternatives are that *c* caused *e* or that it did not, many people would be reluctant to agree with the claim that Jack’s or Jill’s abstention “caused” the Republican victory. Turning to the graded notion, it is also not clear to me whether or not most people would agree that Jack or Jill’s abstentions were even “weak” causes of the Republican victory. It does seem plausible that if “weak cause” is interpreted to mean something like “small causal contribution” many might agree that each individual Republican vote made some such a contribution to the Republican victory and perhaps that Jill’s abstention did so as well. However, for reasons to be discussed below, it is not clear that the normality-based account adequately captures this “causal contribution” idea.

In this connection it is also worth emphasizing that when one talks about the actual causes of election outcomes with large electorates, it is usually substantial blocs of voters who are regarded as candidates for such causes—e.g., Obama to Trump voters in 2016 in states like Wisconsin and Michigan—rather than individual voters. If a shift in a substantial bloc of voters is regarded as, so to speak, a single step in the direction of abnormality or normality (rather than many steps corresponding to changes in individual votes), one can recover, within a framework like Halpern’s, the judgment, endorsed by many, that it was the Obama to Trump votes rather than the votes of those who always vote Republican that caused Trump’s victory or at least that the former was a “stronger” actual cause than the latter. So the normality-based account does seem to capture at least some causal judgments about blocs of voters.

1. **Normality, Causal Selection and Actual Causal Contribution**

I turn now to some more general comments on Halpern’s proposal and others like it. First, to the extent that the proposal captures a set of considerations that influence actual cause judgment, it helps to explain the puzzle raised earlier. It suggests that the reason why it is so hard to write down a rule which takes us from (i) a specification of the laws or structural equations governing a system and (ii) the actual values of the variables characterizing the system to a conclusion about actual cause judgments is that such judgments depend on information in addition to (i) and (ii)—in particular, information about normality. To express the point slightly differently, say that a factor is “intrinsic” to the causal connection between *c* and *e* if and only if it has to do with whether or not *c* and *e* occur or with whether there is a law or causal generalization linking *c* to *e*. Then to the extent that normality considerations play a legitimate role in actual cause judgments, such judgments will incorporate the influence of non-intrinsic (extrinsic) factors. Moreover, to repeat a point made earlier if actual cause judgments incorporate normality considerations, this is a reason (in addition to others described above) to think that such judgments are not a foundation for all other causal judgments, assuming that the latter will not incorporate normality considerations, at least in the same way.

Halpern’s theory aims to cover all actual cause judgments and incorporates a potential role for normality in all such judgements. Assuming that normality tracks selection effects, Halpern’s theory thus seems committed to rejecting the idea that we can distinguish two categories of actual cause judgment, one that reflects selection effects and another, presumably more like Hall’s egalitarian notion (reflecting what I call “actual contribution” below) that does not reflect selection effects. Moreover, regardless of whether Halpern’s theory is accepted, it seems clear from the examples above that many of our intuitive actual cause judgments are influenced by factors associated with selection effects. It may not always be easy to recognize and remove the influence of such factors, in a way that allows us to get at some underlying egalitarian notion just by consulting ordinary intuitive responses. For example, to the extent that we judge that Jill’s abstention was an actual cause of the Republican victory or a stronger such cause than Jack’s abstention, it is arguable that we are making a causal selection judgment, even though this may not be immediately obvious. Thus, even if there is an underlying egalitarian notion that philosophers should reconstruct, reliance on “intuition” or ordinary folk judgment as a way of getting at that notion may be problematic if many intuitive judgments are “infected” by normality type considerations in a way that may not be easy to recognize. At the very least, if there is a more egalitarian notion of actual causation, we need criteria for recognizing when our judgments reflect it and when they do not.

1. **Actual Causal Contribution**

Let me now turn to some remarks exploring this more egalitarian, non-selective notion, beginning with the election scenarios. There seems to be an obvious sense in which each person who voted Republican in Wisconsin in 2016 made exactly the same “actual causal contribution” to the Republican victory. “Actual causal contribution” here is intended to contrast with more generic causal claims that make no reference to actual circumstances or to what actually happened —instead the actual contribution claim is that the votes were causally operative in producing a particular effect (victory) in the specified actual circumstances. This makes it arguable that this actual contribution notion has a serious claim to be considered *a* notion of “actual” causation”, even if it is not the only such notion. To enlarge on this: Assuming the absence of fraud, every Republican vote counts exactly as much as every other vote. The mechanism for aggregating votes is straightforwardly additive, each vote has equal weight, and it is not as though some votes make more of a contribution depending on how others have voted. To be sure, we may also find it natural to describe the Obama to Trump voters as causing Trump’s victory and the votes of regular Republican voters as having a different causal status vis a vis this outcome. However (it might be argued) this seems to involve a practice of causal judgment (one rooted in causal selection considerations) that is different from the actual contribution notion just described. It is not as though the Obama to Trump voters did something that is causally different from the regular Republican voters who voted for Trump. The votes of the former did not have more causal power or “umph” than the votes of the latter—to that extent, thinking of them as having more “causal strength” (as reflected in verbal probes that are interpreted as measuring causal strength) is potentially misleading. Indeed, the reason we find it natural to describe our practice of singling out the Obama to Trump voters as a causal selection practice is exactly because this involves singling out a subset from a larger set of causes (all of the Republican votes), where the latter corresponds to the actual causal contribution notion.

To this we may add the kind of reasoning that we appealed to in support of the original **AC** account of actual causation. Consider first an election between two candidates and just one voter V1. Here no one doubts that V1’s vote for A is an actual cause of A’s victory, that if V1 abstains, this is an actual cause of A’s failure to win and so on. If we assume that as we add voters to the electorate there are no interaction effects of their votes (the votes combine additively) then it seems reasonable to think that each continues to make the same contribution as they would make in the case in which each was the only voter. Thus using an invariance or robustness assumption of the sort described in Section 5 in combination with a symmetry assumption according to which each vote contributes equally leads us from the presence of an actual cause relation in the only one voter case to the presence of such a relation for each voter in the n voter case, even if n voters are far more than are required for a Republican victory.

1. **Features Distinguishing Actual Causal Contribution and Some Additional Illustrations**

I will say more about the possible applicability of this actual contribution notion to other sorts of cases below. But before doing this, it will be worthwhile to explicitly spell out some of the ways in which this notion differs from the notion captured by Halpern’s normality-based account.

9.1. Assuming (as we have throughout) that we are not dealing with cases in which different candidate causes have different probabilistic propensities to produce their effects, the actual contribution notion is not graded and does not come in degrees— either *c* is an actual cause (in the contribution sense) of *e* or it is not. By contrast, as we have seen the normality-based notion does come in degrees.

9.2. To the extent we are dealing with a notion of actual causation that incorporates normality considerations, the associated judgments will be influenced by variations in these. The corresponding claim will not be true to the extent that we are dealing with judgments that involve the actual contribution notion.

9.3. Consider the suggestion made earlier about the (or a) goal or function of actual cause judgments—that these track the effectiveness of possible interventions. This function fits well with a notion that incorporates normality- based considerations. Often an abnormal factor (among all those relevant to some outcome) will be a particular good target for intervention. If one wanted to prevent Trump’s victory in 2016 it would have been a better strategy to try to persuade Obama to Trump voters to change their votes (this would be more likely to be successful) than to target life-long Republicans. By contrast the effectiveness of interventions story fits less comfortably with the actual contribution notion. As the election scenarios illustrate, the actual contribution notion treats each voter symmetrically.

I have used 9.1-3 to contrast normality- based actual cause judgments with a conjectural actual causal contribution notion that behaves somewhat differently. But we can also use 9.1-3 and some of my other remarks above as empirical diagnostic tools. If there is a notion of actual causation in use that looks like the actual contribution notion and not so much like a normality-based notion, perhaps we may be able to use considerations like 9.1-3 to identify instances.

With this possibility in mind, consider again the pre-emption case (EX1.1). If your “intuitive judgments” are like mine, you are likely to regard this as a case in which categorical yes/no rather than a graded judgment seems most natural. That is, the intuitive judgment is that the impact of Suzy’s rock was the actual cause of the shattering and Billy’s rock was not an actual cause at all, rather than an actual cause of less strength than Suzy’s. Note that the “actual causal contribution” notion of actual causation seems to fit this case very well: It was the impact of Suzy’s rock that made an actual causal contribution to the shattering, Billy’s throw made no such contribution. If this is in fact how people judge, one might expect that when presented with such a scenario, subjects would assign a very high causal rating to Suzy’s throw (a rating at ceiling or nearly so) and a very low rating to Billy’s throw, where this is interpretable as “not a cause at all”). In contrast, if Billy’s throw were to receive some more intermediate score, this would indicate that, contrary to what I have suggested, subjects are willing to make graded judgments in this sort of case. As far as I know, examples of this sort have not been studied empirically. Such investigations might help to determine whether there are cases in which people judge in accord with an ungraded causal contribution notion rather than a more graded notion.

Next consider whether normality manipulations affect actual cause judgments in various scenarios. Suppose it is extremely common and not at all statistically abnormal for Suzy to throw rocks at bottles and to shatter them and extremely rare (abnormal) for Billy to do this, although both are equally efficacious when they do throw. Or suppose that the rules permit Suzy to shatter bottles but prohibit Billy from doing so. Or suppose Suzy and Billy have the opposite abnormality/ normality profiles. Would variation of any of these factors make a difference for most people’s actual cause judgments in a pre-emption case concerning which events involving rock impacts were the actual causes of a bottle shattering? Although there is some contrary evidence from somewhat similar experiments[[13]](#footnote-13), my guess is that such normality manipulations would have little effect on people’s actual cause judgments in scenarios of the sort described involving collisions and event causation. If so, this again may be evidence that we are dealing with a kind of actual cause judgment that is somewhat different than the notion involved in normality-based selection judgments.

Finally consider the effectiveness of intervention account of the function of actual cause judgments. We noted above that this seems to fit many actual cause judgments that are influenced by normality considerations/causal selection rather well. By contrast this account provides a less satisfying treatment of the pre-emption example. If our goal is to identify the intervention that would be most effective in preventing the shattering of the vase, it isn’t clear why we should distinguish so sharply between Suzy’s and Billy’s throw. Given the stipulations of the example, even if we had prevented Suzy from throwing or deflected her rock, the vase still would have shattered because of Billy’s throw. Even if we think that in any realistic case, there would be some non-trivial probability that Billy would have missed, had Suzy not shattered the vase, one would expect, on the effectiveness of intervention account, for this to result in graded judgments of actual causation, with a slightly higher causal strength judgment for Suzy than for Billy. (We know that Suzy’s rock hit the vase, so that preventing her throw while not preventing Billy’s would replace a situation in which shattering occurs with one in which it is merely probable that shattering occurs.) In this connection, we may note that asking subjects how effective they think various interventions would be might be another way of tracking whether we are dealing with different varieties of actual cause judgments: if subjects agree that preventing Suzy’s throw would not be a particularly effective way of preventing bottle shattering, while agreeing that Suzy’s throw was an actual cause, this would suggest that the effectiveness of intervention rationale does not capture their judgments .

1. **Some Methodological Reflections**

The possibility just floated—that there may be at least two different varieties of actual cause judgments—raises several methodological issues. One obvious point is that to the extent this is the case, it is presumably a mistake to look for a single theory of actual causation that covers both kinds of judgment. Indeed, if we are dealing with two different kinds of judgments with sensitivities to different factors, this might help to explain why finding a single theory that covers everything that we may think of as an actual cause judgment has turned out to be so difficult.

However, if we take the route of supposing that there are several different sorts of actual cause judgments, requiring different theories, we face the issue of how to recognize cases of each sort of judgment. The criteria 9.1-3 above may help with this but as also suggested in Section 8 more fine-grained verbal probes (and use of a variety of these rather than a single probe) may prove fruitful. One motivation for this is that standard causal strength questions track or are sensitive to a number of distinct and apparently independent dimensions of causal assessment. For example, type-level causal relations that are more stable or robust in the sense of Woodward (2006) tend to receive higher strength judgments (Vasilyeva et al., 2018) and when causal relations are probabilistic, causes that boost the probability of an effect more highly receive higher strength ratings (e.g. Cheng, 1997). All this is in addition to the fact that strength judgments are affected by normality considerations (which presumably are at least somewhat independent of robustness and probability raising considerations) in the case of actual cause judgments. Probes that are more sensitive to differences among actual cause judgments and between these and other sorts of judgments might help to ameliorate this limitation. For example, rather than only providing subjects with different actual cause claims to be rated according to strength, researchers might also describe scenarios and (i) ask subjects what “the” cause or causes of some effect are in the scenarios or, alternatively (ii) ask them to list “all” of the causes they can think of or anything that qualifies as “a cause”. (In other words, subjects are asked to generate their own lists of causes rather than assessing causal claims provided by the researcher.) Possibly, if there are differences in answers to (i) and (ii) this might suggest that (i) tracks a selection-influenced notion of causal selection while (ii) tracks a notion that is less sensitive to selective considerations. As an illustration, suppose that in the pen case, when asked for “the” cause of the secretary’s being unable to take the message, subjects tend to select the professor’s taking the pen but when asked to list all of the causes, they also include the action of the staff person, so that the answers to (i) and (ii) are different. By contrast, suppose that in a pre-emption scenario like EX5.1, subjects give the same answer to (i) and (ii) – in both cases only the pre-empting cause is cited. This might be taken to reflect a difference between what the subjects are doing in the two sorts of cases, a difference that might be less apparent if a generic causal strength probe is employed. More generally, given that there are many different verbal probes that might be employed, it would be very desirable to understand how they relate to one another—when they give concordant answers and when they do not. For example, when subjects give a higher causal strength rating (according to some standard probe) to (i) *c* caused *e* than to (ii) *c\** caused *e* in some scenario, do they spontaneously cite (i) rather than (ii) when asked about “the cause” of *e* in the scenario? [[14]](#footnote-14)

As another possibility consider the target of intervention idea as an account of the function of actual cause judgment. If this is correct (or to the extent it is correct) one would expect that when subjects are asked what they would do to change some outcome in a scenario, the results would mirror their actual cause judgments as measured by some other verbal probe. If one were to find that in some cases actual cause judgments correlate closely with targets of intervention queries and in other cases do not, this again might suggest that we are dealing with two different kinds of judgments. For example, actual cause judgments in the election scenarios considered above might track target of intervention questions well but this might not be the case for scenarios like (EX5.1).

1. **What Might be the Function of the Actual Contribution Notion?**

I suggested above that the target of intervention notion provides a fairly intuitive account of the function of some actual cause judgments in which selection seems to be operative but I have not suggested a function for the actual contribution notion (supposing that there is such a thing). What work does the latter do? One possibility is that the robustness features of this notion make it better adapted to generalization or transportability to other contexts than a notion in which selection figures more centrally. In particular, suppose that we think of “actual contribution” in the following way: If *c* makes such and such an actual causal contribution x to *e* in one set of circumstances then insofar as *c* makes an actual contribution to *e* in other circumstances, we expect it to make the same contribution x, at least for a range of such circumstances, so that a kind of invariance or robustness of this contribution is present. This is the sense in which a vote for a candidate makes the same actual causal contribution regardless of how others vote. Similarly, when asked whether a drug is an actual cause of some effect in a patient, we tend to frame this as an actual contribution question because we proceed on the assumption that if the drug is efficacious at all, it will be an actual cause of broadly the same effect in many others. (We think of a vaccine as making the same contribution to immunity in different people, regardless of whether administration of the vaccine is “normal.”) By contrast since selection effects are sensitive to such “extrinsic” factors as the frequency with which a cause occurs in a population or how likely it is that a cause will assume some alternative value if it does not take its actual value, causal selection judgments will tend not to exhibit this sort of robustness. Since generalizability is among the considerations we care about in making causal judgments, this suggests one kind of rationale or function for actual causal contribution judgments—a rationale that seems somewhat independent of the target of intervention considerations that influence causal selection.

To this we may add the observation that to the extent that responsibility and credit and blame judgments are influenced by causal judgments, the actual contribution notion seems to play an important role (of course as illustrated earlier, normality-based considerations can also matter). In the overdetermination case EX5.2 we hold both shooters responsible for the death because we think of them as making the same actual contribution to the death. In the rock throwing example, Suzy’s throw and not Billy’s is blamed or credited with the destruction of the bottle because it is her throw that made the actual contribution. Billy is guilty of something else—e.g., attempted bottle destruction.

1. **Conclusion**

If the arguments in this chapter are defensible then, whatever else might be said about actual causation, it turns out to be complicated. There appears to be some reason to think that actual causation is somewhat heterogenous, incorporating at least two different notions or elements, one more closely linked to causal selection and the other to a notion that tracks something more like actual contribution. But in some cases the line between these two notions is rather fuzzy as a number of the examples above illustrate. Given the complexities of the notion, it seems a bad strategy to take actual causation (at least as this is reflected in our intuitive judgments) as metaphysically fundamental or as a kind of stand-in for causation in general.

**13. Appendix**

**13.1. Proximate vs Distal**. A form of actual cause judgment that is quite widespread but which has received relatively little attention from philosophers, psychologists or formal modelers involves selecting among more or less distal or proximate causes. In such cases there are (often complex) chains or webs of causes leading up to an outcome but only some of these will be described as actual causes of the outcome-- or at least some will be preferentially selected as "actual causes". Judgments about causes of death provide one set of examples. When a patient’s heart stops beating as a consequence of a very severe covid-19 infection the actual cause of the patient’s death may be judged to be (or may be attributed to) the infection despite the fact that the heart stoppage was the more immediate or proximal cause of death. Similar judgments are made when an AIDs patient dies of pneumonia and in many other cases. Judgments of this sort are presumably influenced in part by assessments that the selected cause (e.g., Covid) made the death nearly inevitable, although it might have caused the death by any one of several different routes. Such judgments also make sense within an efficiency of intervention account of the function of actual cause judgments, since preventing the Covid infection would have been the most effective way of preventing death. (Given the infection, if the patient had not died of heart stoppage, she likely would have died in some other way.)

At the same time there appears to be a considerable conventional element to such judgments, with criteria for cause of death varying widely across different countries and jurisdictions and reflecting choices that look quite arbitrary. For example, in some jurisdictions the rule is that when patients undergo surgery for advanced cancer and die within 30 days of the surgery, their deaths are attributed to the surgery; if they die later the cancer is judged as the cause of their deaths. Other jurisdictions have different rules [[15]](#footnote-15). Despite the fact that such practices seem to involve causal selection, it is not clear that “normality” considerations give us a good handle on them, unless “normality” means nothing more than “in accordance with convention”, although (to anticipate **13. 2** ) judgments of inevitability or time order information may also play a role . Accounts of actual cause judgment that do not make use of normality considerations such as **AC** (Section **5)** also do not capture the distinctions made in such judgments.

The criteria for the actual causes of death that are widely in use have another noteworthy feature. Although they sometime seem to track effectiveness of intervention criteria, in other cases they do not. It is clear that in many cases the most robust causes of death and disability involving disease are under “social” control – they include such distal causes as malnutrition, environmental pollution, various forms of stress that suppress immune function and so on. (These causes are “robust” in the sense that given their presence, premature death is likely to result via one route or another.) These causes are also the most effective targets of intervention if one wants to reduce the incidence of death and disability. Yet in most cases they are not on the list of acceptable candidates for actual causes of death that are cited in an autopsy or coroner’s report. Their omission presumably reflects both the professional interests and competencies of medical personnel and ideological factors. A similar observation holds for many claims about the actual causes of industrial accidents, where such claims may cite failures of individuals to operate equipment adequately rather than the conditions that make it possible for such failures to occur and have catastrophic consequences. An effectiveness of intervention rationale for actual claims would thus regard many of the judgments people make about such cases as defective, even if they are widely shared[[16]](#footnote-16). As these examples illustrate, although actual cause judgments can be useful, they can also be about larger patterns of generic causal dependence[[17]](#footnote-17).

**13. 2**. Processes in which outcomes become virtually inevitable at earlier times and candidate causes occur at later times present another source of difficulty for standard accounts of actual causation. As noted in **13.1** disease causation provides one set of examples, but consider the following:

Ex 13.2. 1. A virus is widespread. Some people are constitutionally immune—it is certain they will not get the virus if exposed. A vaccine is also available that is 100 per cent effective in preventing infection. Jones is constitutionally immune but does not know this and so takes the vaccine. Was his taking the vaccine an actual cause of his not becoming infected? I assume that many people will say “no”— whether or not he took vaccine, Jones was not going to get infected. The appropriate causal model of this case will have an arrow from Jones' immunity *I* to his failure *F* to be infected and no arrow from *V* (his taking the vaccine) to *F*. But why not instead think of the case as as isomorphic to EX 5.2, with *I* and *V* as independent sufficient causes of *F*, so that the example involves overdetermination?

EX 13.2. 2. The Pagans are playing the Christians. They lead 42-0 until the final 30 seconds when Thor scores another touchdown for the Pagans. Was Thor’s scoring an actual cause of the Pagan’s victory? **AC** says “yes” but it isn’t clear that this would be the confident judgment of most people. A normality based treatment like Halpern's might judge that Thor's scoring was a weak cause of the victory, depending on how the notion of "normality" is understood. (See below.) This judgment may be defensible but the normality account in its present formulation seems to omit a consideration that underlies that judgment. This has to do with the role of time.

In both EX 13.2.1-2 a cause or causes of the final effect are present at an earlier time and make the effect inevitable or effectively so. This encourages us to judge that an additional candidate later cause was not efficacious in contributing to the effect. The role of time seems crucial since this seems to be what leads us to judge that the effect was “baked in” by earlier occurrences. However there is no explicit role for temporal order in **AC** or for that matter in the normality framework. One possible strategy for dealing with EX13.2.1-2 would be to employ a normality ordering to which temporal constraints and perhaps considerations reflecting inevitability/immutability are added— contingencies that involve changing the states of off-path causes that occur earlier might be viewed as larger departures from normality than changes that occur later and changes in characteristics that are seen as fixed and immutable (like immune status) might be viewed as very large departures from normality.

This of course leads again to a graded notion of actual cause, something that is perhaps not unreasonable in the case of EX13.2.2 but perhaps less reasonable for EX13.2.1, where one might think that *V* is just not a cause of *F*.

Regarding EX13.2.2 we can add that it provides an additional illustration of the distinction between the selective and non-selective notion of actual causation: in one sense, Thor’s touchdown makes exactly the same contribution to the Pagans’ victory as each of other touchdowns, in another sense, based on selective considerations, it seems less of a cause.

Finally consider two variants on EX 13.2.2. In the first, the score is Pagans 40, Christians 38 at the beginning of the fourth quarter. Thor scores early in this quarter so that the final score is 47 to 38. At the time Thor scores, the Pagans’ victory seems less than inevitable. Does this make a difference to our actual cause judgment? Is Thor's score a "stronger" cause in this case than in the 42-0 case even though, given what actually happens, the Pagans would have won without Thor's score?

In the second variant the game is called off in the middle of the fourth quarter because of weather with the rules dictating that the team that is ahead wins. Was the weather the actual cause of the Pagans’ victory (a pre-empting cause)[[18]](#footnote-18)? Does it matter if the score was 42 to 0 or 40 to 38 at that point? I suspect that actual cause judgment gets blurry in some of these cases—our practices may not dictate clear answers and it is also unclear why they *should* dictate such answers. These are additional reasons for thinking that actual cause judgments are not foundational in the way that some metaphysicians have supposed.

**References**

Bogen, J. and Woodward, J. (1988) “Saving the Phenomena” *Philosophical Review* 97: 303-352.

Cheng, P.W. (1997). “From Covariation to Causation: A Causal Power Theory” Psychological Review, 104,367-405.

Knobe, J. & Fraser, B. (2008). “Causal Judgment and Moral Judgment: Two Experiments”

In W. Sinnott-Armstrong, *Moral Psychology*, Cambridge, MA: MIT Press. 441-448.

Glymour, C., D. Danks, B. Glymour, F. Eberhardt, J. Ramsey, R. Scheines, P. Spirtes, C.M. Teng, and J. Zhang (2010). “Actual causation: a stone soup essay” . *Synthese* 175: 169–192.

Hall, N. (2004) “Two Concepts of Causation”, in *Counterfactuals and Causation*, edited by John Collins, Ned Hall, and L. A. Paul, MIT Press, Cambridge MA, pp. 225-276.

Halpern, J. and Hitchcock, C. (2015) “Graded Causation and Defaults” *British Journal for the Philosophy of Science* 66: 413-457.

Gerstenberg, T., Halpern, J. and Tenebaum, J. (2015) “Responsibility Judgments in Voting Scenarios” web.mit.edu.tger/www/papers.

Glymour, C., D. Danks, B. Glymour, F. Eberhardt, J. Ramsey, R. Scheines, P. Spirtes, C.M. Teng, and J. Zhang (2010). “Actual causation: a stone soup essay” . *Synthese* 175, 169–192.

Halpern, J. (2016) *Actual Causality*. Cambridge: MIT Press.

Halpern, J. (2018) “Appropriate Causal Models and the Stability of Causation” <https://arxiv.org/abs/1412.3518v2> asXiv:1412.3518. pdf

Halpern, J. and Pearl, J. (2005). “Causes and explanations: A structural-model approach. Part I: Causes”. *British Journal for Philosophy of Science* 56: 843–887.

Hitchcock, C. (2001). “The intransitivity of causation revealed in equations and graphs”. *Journal of Philosophy* XCVIII: 273–299.

Hitchcock, C. (2007). “Prevention, preemption, and the principle of sufficient reason”. *Philosophical Review* 116:495–532.

Hitchcock, C. and Knobe, J. (2009) “Cause and Norm” *Journal of Philosophy* 106: 587-612.

Icard, T., Kominsky, J., & Knobe, J. “Normality and Actual Causal Strength” *Cognition* 161:80-93.

Knobe, J and Fraser, B. (2008) Knobe, “Causal Judgment and Moral Judgment: Two Experiments” In W. Sinnott-Armstrong, *Moral Psychology*, Cambridge, MA: MIT Press. 441-448.

Kominsky, J., Phillips, J., Gerstenberg, T., Lagnado, D. and Knobe, J. (2015).”Causal Superseding”  Cognition, 137, 196-209.

Livengood, J. (2013). “Actual causation in simple voting scenarios”. *Nous* 47: 316–345.

Northcott, R. (Unpublished) “Pre-emption cases may support, not undermine, the counterfactual theory of causation” Philsci Archive.

Paul, L. and Hall, N. (2013.) *Causation: A Users Guide*. Oxford University Press.

Sytsma, J. and Livengood, J. (2020) “Actual Causation and Compositionality” *Philosophy of Science* 87: 43-69.

Samland, J. and Waldman, M. (2016) “How prescriptive norms influence causal inferences” *Congnition* 156: 164-176

Strevens, M. (2008) *Depth: An Account of Scientific Explanation*. Cambridge Harvard University Press.

Vasilyeva, N., Blanchard, T., and Lombrozo, T. (2018). "Stable causal relationships are

better causal relationships." *Cognitive Science,* 42(4), 1265-1296.

Woodward, J. (2003). *Making Things Happen: A Theory of Causal Explanation*. New York: Oxford University Press.

Woodward, J. (2006) " Sensitive and Insensitive Causation”. *The Philosophical Review*.115: 1-50.

Woodward, J. (2014) “A Functional Account of Causation” *Philosophy of Science* 81*:* 691-713.

Woodward, J. (2020) “Causal Attribution, Counterfactuals, and Disease Interventions” In Eyal, Hurst, Murray, Schroeder, and Wikler (eds) *Measuring the Global Burden of Disease: Philosophical Dimensions.* Oxford: Oxford University Press

Woodward, J. (2021) *Causation with a Human Face: Normative Theory and Descriptive Psychology.* New York: Oxford University Press.

1. The terminology of “actual causation” seems to be becoming increasingly standard. In the not very recent past in philosophy the judgments in question were commonly described as involving “token causation”, as opposed to “type causation”—see below for discussion of this terminology. Another term that is sometimes employed to characterize the claims/judgments in question is “singular causation”. [↑](#footnote-ref-1)
2. Woodward, 2021, discusses a number of additional examples in which the wording of the verbal probe used in eliciting causal judgments makes a difference to the reported results. Other papers (among many) showing how variation in wording affects results include Samland and Waldman (2016) and Sytsma and Livengood (2020). [↑](#footnote-ref-2)
3. Put differently, we need to distinguish what the having of the intuition is evidence for from whether the claim reported in the intuition is true. The having of the intuition can be evidence for how people judge but the having per se is not evidence that the claim reported in the intuition is true—the latter has to do with such worldly facts as the ability of rocks to break bottles, the fact that Suzy’s rock came in contact with the bottle and so on. If matters are as reported in the scenario, it is likely that the reported causal claim is true but that is not because having the intuition is a way of establishing that it is true. [↑](#footnote-ref-3)
4. For discussion of a number of examples in which this has turned out to be the case, see Woodward forthcoming. [↑](#footnote-ref-4)
5. Again, see Woodward 2021 for examples. [↑](#footnote-ref-5)
6. In a disjunctive causal system multiple events are each sufficient to produce an outcome. “Abnormal deflation” refers to the tendency of people to judge such events in disjunctive systems as more causal to the extent that they are less rare. [↑](#footnote-ref-6)
7. This is often implicit but it is hard to make sense of the nearly exclusive focus on actual cause judgments in writers like Paul and Hall, 2013 unless they are assuming that there is something fundamental about such claims or that the interesting features of other kinds of causal claims are already present in actual cause claims. Strevens (2008) is one writer who explicitly takes actual causation as his starting point. [↑](#footnote-ref-7)
8. For example, this is so for the Cheng’s (1997) causal power theory. Woodward (forthcoming) argues that is a deep feature of many causal generalizations and laws that they are expected to be invariant across change in the frequency with which the independent variables are instantiated. [↑](#footnote-ref-8)
9. The impact of normality on causal judgment will depend both on (i) the way in which the scenario is modeled and (ii) the details of the account of actual causation adopted and on the way in which normality figures in that account (see below). However, as an empirical matter, normality has some influence on causal selection in both conjunctive and disjunctive structures and formal theories like those do to Icard et al. (2017) and Halpern (2016, 2018) predict such influences. So if it is correct to assimilate the rock throwing example to other cases of causal selection one would expect some influence of normality on the former. [↑](#footnote-ref-9)
10. Another way of putting this issue is in terms of function. Full information about the equations or dependency relations governing a system and the actual values taken by variables enables one to answer a range R of counterfactual questions about how the system would behave under various settings of the variables. If this information is not enough to fully answer queries about the actual cause relations in the system, such queries must correspond to different questions than those in R. This suggests that the point or function of actual cause claims is not to answer questions like those in R. [↑](#footnote-ref-10)
11. Halpern has confirmed this assessment in recent correspondence. [↑](#footnote-ref-11)
12. I am grateful to Alex Wiegmann for drawing my attention to Gerstenberg et al. 2015 which is reports the results of a study of ordinary subject judgments of *responsibility* for various election outcomes. Assuming that subject’s actual cause judgments are related to (although perhaps not identical with) their responsibility judgments, it is arguable this paper provides information about the former. In the election scenarios studied by Gerstenberg et al. a number of distinguishable factors seem to influence responsibility judgments—these include the extent to which a vote is pivotal, and what the authors call dispositional and situational normality. Perhaps one might think of all of these as having to do with normality in some extended sense, but if so, normality will be a complex and apparently multidimensional notion. [↑](#footnote-ref-12)
13. Icard et al. (2017) report evidence of what they call “abnormal deflation” according to which if *E* depends disjunctively on both *A* and *C*, both occuring, people will be more inclinded to say that *C* caused *E* if *C* is normal rather than abnormal. These of course are overdetermination rather than pre-emption scenarios . Moreover, the verbal probe asks about the causal role of agents (Billy and Suzy) rather than events, which may well make an important difference, as shown in Samland and Waldmann (2016)—questions about agents seem more likely to be interpreted as questions about blame. In addition, the scenarios themselves did not involve collision events such as rock throwing. Still I acknowledge that the results are surprising: if a similar pattern holds in a scenario in which two simultaneous impacts are each sufficient for bottle shattering, people will assign the impact of Suzy’s rock a higher causal strength rating than Billy’s to the extent that Suzy’s rock throwing is more normal – a pattern of judgment which I at least would not have expected. In any case one wonders whether similar effects would be present in pre-emption scenarios. Interestingly the normality or abnormality of *A* in disjunctive structures does not seem to affect judgments about the causal status of *C*. [↑](#footnote-ref-13)
14. These issues about different verbal probes and their interrelation also have implications for the common philosophical practice of appealing to intuitions about cases. When different philosophers do this they may be asking themselves different questions or failing to recognize ambiguities in the questions that they ask themselves. If so, the “same” case may elicit different responses for reasons that are unrecognized. Again, see Woodward (2021) [↑](#footnote-ref-14)
15. See Woodward, 2020 for details. [↑](#footnote-ref-15)
16. I don't take this to be an objection to the effectiveness of intervention rationale construed as a normative proposal. It rather illustrates how thinking about the function of actual cause judgment can have a normative bite and involves more than just fitting the actual cause judgments people are disposed to make. [↑](#footnote-ref-16)
17. See Woodward, 2020 for more on this theme. There are many circumstances in which information about the actual causes of death of the sort reported in autopsies will tell us little about the most effective measures for preventing death in that population. [↑](#footnote-ref-17)
18. This is a variant on a case suggested by Northcott (unpublished) [↑](#footnote-ref-18)