LSE CPNSS

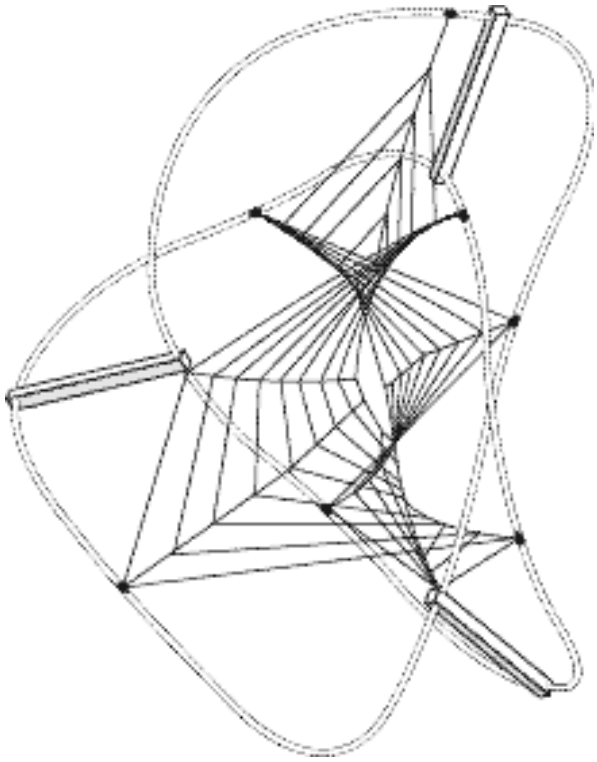
**Centre for Philosophy of Natural and Social Science**

**Causality: Metaphysics and Methods**

Technical Report 08/03

# *How Can We Know What Made the Ratman Sick? Singular Causes and Population Probabilities. An essay in Honour of Adolf Grunbaum*

Nancy Cartwright



Editor: Julian Reiss

# How Can We Know What Made the Ratman Sick?
# Singular Causes and Population Probabilities
# An Essay in Honour of Adolf Grunbaum*

Nancy Cartwright
Department of Philosophy, Logic and Scientific Method
London School of Economics
and
Department of Philosophy
University of California-San Diego

October 2002

*Research for this paper was supported by a grant from the Latsis Foundation and by a grant from UCSD that provided research assistance. I am grateful for both. I first became seized with the importance and excitement of philosophy of science from Adolf Grunbaum as well as with a respect for careful thought and serious dialogue, which I am afraid I have never been able to emulate sufficiently; and I am deeply grateful to him, as well as for his support and friendship over many years.

Much of Adolf Grunbaum's work on psychoanalytic theory over the years has been concerned with causal claims. These can be claims either about the aetiology of psychic disorders or about the efficacy of various aspects of psychoanalytic treatment; and they can be either singular causal claims intended to describe a single individual, such as "Anna O's neurotic symptoms were cured by Breuer's inducement of cathartic recall"[1], or generic claims meant to be true in general (what I call 'causal laws'), like "The adult male disposition of castration anxiety is attributable to oedipal childhood events prior to age 6".[2]

I am here going to discuss singular causal claims and how to support them. Grunbaum and I both[3], I gather, accept what is conventionally regarded as best practice in so-called 'quasi-experiments' – so long as all the precautions are tended to. I have in mind here methods like pretest-postest control group designs, Solomon four groups designs, simple ANOVA or Latin square designs.

In all these cases we look for a difference in the outcome between when the individual is "treated" with the cause and when not. The methods in my list are increasingly difficult to apply, and correlatively, increasingly more powerful. The simpler designs are more apt to go astray because they are less good at guarding against reasons for the difference in outcome *other than* the treatment. This is the same idea that Grunbaum uses in his criticisms of many of the psychoanalytic claims about individual aetiologies: the reasons employed in arriving at a causal conclusion are not good enough to ensure that the putative cause made a difference to the outcome.

I do not agree with Grunbaums's insistence that the cause must make a difference in the single case. The cause must *contribute* to the effect. But that is no guarantee that it will make a difference. The idea that it must make a difference lies at the core of counterfactual accounts of causality; and these notoriously give wrong verdicts in cases of preemption. The President would nevertheless have died if the trainee assassin had not fired because in that case the experienced assassin standing beside the trainee would have shot.

There are of course ways to adjust our semantics for either counterfactuals or causal claims to deal with cases of preemption. For instance, withdraw the claim that causes make a difference to their effects and replace it by a two-stage analysis. One factor causes another if there is a chain of causal dependence between the cause and the effect, where the counterfactual connection does hold for each link. We may dispute whether this tactic does the job since it depends on the empirical hypothesis that between a cause and its effect there is always a chain of events where each is so close – 'close enough' – to the next that preemption is impossible and the first will make a difference to the second. But whether it works or not, it does not rescue the claim that a cause must make a difference to its effect but rather gives up on it. I maintain that that is all to the good. For the claim mistakes a 'sometimes' test, or symptom, of causation for a necessary condition: 'sometimes', perhaps often, causes contribute to their

[1] Grunbaum 1993, pp. 238ff.

[2] *ibid.*, p. 176

[3] *Cf.* Grunbaum's endorsement that "the A-B-A does succeed in discrediting spontaneous remission", *ibid.*, p. 237.

effects in such a way as to make a difference. But it is not necessary that they do so.

Even should we adopt the assumption that causes make a difference we still face enormous difficulties in putting it to use. That will be my topic here. How can we tell whether one factor makes a difference or not to another, outside of one of the really good single case designs? Grunbaum offers a necessary condition, which I will describe in a moment. A necessary condition can of course rule out singular causal claims, even though it is not enough to admit them. Here I am more cautious even than Grunbaum. I do not know of any general criteria I am happy with that will rule out singular claims as false, let alone some criteria that will rule them in as true. Outside of a good quasi-experiment, unless we are in a situation where we have a great deal of background knowledge to deploy and are not relying on general criteria (*i.e.*, on ones that are valid in every case),[4] the verdict must be left open.

The problem of 'other reasons' for a positive outcome is very familiar in the case of treatments. The patient got better, but was it the treatment that did it as opposed to 'spontaneous remission'? Lots of conditions just do go away 'on their own'. Or might it instead be a 'placebo effect' – the result of confidence and expectation that does not depend on the specific character of the treatment but might accompany any treatment?[5] These are just the problems Grunbaum raises. For instance, Anna O, it seemed, lost her neurotic symptoms under Breuer's treatment of cathartic recall. But what convincing reason is there to think the treatment caused the relief?[6]

Similarly for claims about the sources of an individual patient's neuroses. Was the Ratman's obsession provoked by his encounter with the Czech captain? Was its base cause certain repressed (or imagined) childhood experiences? Was Anna O's inability to drink water due to her disgust at seeing a dog lapping water from a friend's glass? Much of the argument in favour of these hypotheses depends, Grunbaum notes, on "thematic affinities"; and thematic affinities are not enough!

In explaining why thematic affinities are not enough, Grunbaum considers a couple of analogies. He asks what licenses a tourist to make the causal inference that shapes in the sand that look very much like human footprints were caused by human feet walking there. He says

> To draw the inference, the tourist avails himself of a crucial piece of additional information [...] footlike beach formations in the sand never or hardly ever result from the mere collocation of sand particles under the action of air, such as gusts of wind. Indeed, the

[4] I do not mean here that the criterion can actually be applied in every case, but rather that its deliverances are reliable universally if properly applied.

[5] To be more careful, we should follow Grunbaum's definition: "I speak of a treatment gain as a *placebo effect* with respect to a particular target disorder, therapeutic *theory*, and type of patient, just when that positive effect is produced by treatment factors *other than* those designated as the efficacious ones by the therapeutic theory", *ibid.*, p. 189.

[6] *Cf. ibid.,* pp. 237ff.

> additional evidence is that, within the class of beaches, the incursion of a pedestrian onto the beach *makes the difference* between the absence and presence of the footlike formations.[7]

He also contrasts two dreams about houses. The first is Agnes's dream of a house that looks just like Frank Lloyd Wright's *Falling Water*. She has the dream just after she has visited it for the first time, never having seen or heard about it before then. The second is his own dream about houses after a typical day on which he has passed a lot of houses. In the first we have reason to postulate a causal connection because "Agnes's visit *made a difference* to her having that dream."[8] The contrary is the case with Grunbaum's own dream:

> To put it more precisely, seeing a house on the day before a dream does not divide the class of the day's waking experiences on the prior days into two subclasses, such that the probabilities (or frequencies) of the appearance of a house in the next dream *differ* as between the two subclasses. On the other hand, in Agnes's life, such a division does occur, with ensuing *different* probabilities of dreaming about that house.[9]

More generally:

> If $X$ is to be causally relevant to $Y$ in reference class $C$, $X$ must *partition* $C$ into *two* subclasses in which the probabilities or incidences of $Y$ are different from one another.[10]

He repeats this general demand *verbatim* in his 1992 "Postscript" to his criticism of the way the case study method is used in psychoanalysis in Grunbaum 1993; and he notes there that this criterion is defended by Wesley Salmon in his work on statistical relevance as a mark of causation. It is this condition I want to focus on.

The first thing I want to note is that this is, as I said, offered as a necessary condition. Showing that $X$ increases the probability of $Y$ in the right reference class $C$ shows that $X$ can cause $Y$; on my account of causal laws[11] it even shows that some individual $X$'s do cause $Y$'s in $C$. But that does not show that for the case at hand $X$ did cause $Y$. A lot of other factors that can cause $Y$ might be present as well and they might have been responsible for $Y$. Or $Y$ might have, in a sense, had no cause at all, as in the examples of spontaneous remission – $Y$'s occurrence was the outcome of the natural development of the system. The requisite increase in probability of $Y$ on $X$ is not enough to show that Grunbaum's own condition – that $X$ made a difference to $Y$ in the case at hand – is met.

[7] Grunbaum 1990, p. 568
[8] *ibid.*, p. 570
[9] *ibid.*
[10] *ibid.*, p. 571
[11] *Cf.* Cartwright 1989.

This might be overlooked if we focus only on the two analogues Grunbaum mentions. For in each case the other kinds of factors that raise the probability of the effect are extremely unlikely to have occurred. For instance, the prints in the sand could be made by a hand-held mould as part of an elaborate hoax, or a line of children in a strange sand castle competition. If these had occurred, then more needs to be done to show that the putative cause is really the one that brought about the effect (or, where the effect is cumulative, that it contributed).

Let us turn now to the condition itself. I have been reading this in the usual way, as a condition intended to bear immediately on *generic* causal claims, not singular ones, on what I call "causal laws". My paradigms are "Aspirins relieve headaches" and "Inverting and sparking a population of atoms causes lasing". For generic claims, I want to argue, the condition may be a sufficient condition, but it is not necessary.

Whether it will be a sufficient condition in a reference class depends very much on what that reference class is like. We are all familiar with the problem of spurious correlation. One factor raises the probability of a second, but it does not cause it. Rather, they are both joint effects of a common cause. A common tactic here is to demand reference classes in which all common causes are held fixed. But that is not enough. Even in a reference class like that, if the members of the class are also preselected for a given effect, causes of that effect can be correlated or anti-correlated, without causing each other. Also, if causes can act purely probabilistically, then joint of effects of a common cause can be correlated even in a population where the cause is held fixed. There are also those strange quantum mechanical cases, where distant factors are correlated even – so the standard story goes – without any causal explanation.

I think the best we can do here is to employ a process of elimination. Causes can increase the probability of their effects. But there are also many other possible explanations for such an increase in probability. If we have a reference class in which all other explanations for the increase in probability have been eliminated, then it must be due to the fact that the one factor causes the other.

Next, why do I maintain that the condition is not necessary? There are two well-known reasons. For the first, let us consider a version of Wesley Salmon's example,[12] which was central to his work on statistical relevance that Grunbaum cites. We consider an experimental situation in which by the random flip of a fair coin either a strong radioactive element, $E_1$, or a weak one, $E_2$, will be introduced into an empty box. We look for the presence of a radioactive particle, $\alpha$, later. The probability that $E_1$ produces an $\alpha$ particle in the designated time is .9; for $E_2$, the probability of $\alpha$ is .1. Since the probability is .5 for either element to be in the box, the overall probability for $\alpha$ is .5. Notice that the probability of $\alpha$ with $E_2$ in the box is lower than this. But when $E_2$ is in the box and an $\alpha$ particle appears, it is certainly $E_2$ that causes it. This led Salmon to claim that causes may reduce the probability of their effects as well as raising it.

A slight change in the example leads to a more radical conclusion, however.

[12] Salmon 1971

Imagine we have available a large number of different radioactive elements; two – $E_3$ and $E_4$ – happen to have the same probability for producing $\alpha$ particles. Repeat the experiment now with $E_3$ and $E_4$. If $E_3$ is in the box when an $\alpha$ particle appears, definitely $E_3$ causes it and similarly with $E_4$. Yet in this case the probability stays the same. The cause makes no difference to the probability of the effect.

Salmon's example is a case of Simpson's paradox, which I rediscovered years ago in criticizing Patrick Suppes's use of increase in probability as a necessary condition for causality.[13] Suppose we partition a population into cells according to different values of a variable $Z$. Then $X$ can increase the probability of $Y$ in every cell of the partition and yet not do so in the population as a whole. In order for this to happen $X$ and $Y$ must be probabilistically dependent on $Z$. In my amended version of Salmon's example $X$ = radioactive element 1, $Y$ = presence of decay particle, and $Z$ = presence of radioactive element 2. $X$ increases the probability of $Y$ given $Z$ and it also increases it given $-Z$; but it does not do so in the population as a whole. That is because, as we have seen, the presence of the one element is correlated with the absence of the other.

One proposal conventionally offered in solution to this problem is the one I described above in talking about sufficient conditions for causal claims: insist that the reference class for testing "$X$ causes $Y$" should be homogeneous with respect to all other causes of $Y$ other than $X$. In a reference class like that, nothing should be correlated (or, so it is supposed) with $X$, and hence Simpson's paradox cannot make problems. I do not know what Grunbaum thinks of this solution. He might have intended this as a constraint on the reference class $C$ all along. Probably he does discuss this somewhere that I missed. Here is a remark I did find, in a discussion of a similar problem in a singular case:

> But it is *utterly chimerical* to predicate a research design on a situation in which two people differ *only* with respect to the property that the investigator *conjecturally* deems relevant to the outcome![14]

Of course the quote only says that we will not be likely to ensure such a research design; it does not tell us whether our condition is a necessary condition if we did have such a design.

Besides the strategy of demanding an increase in probability only in reference classes that are homogeneous with respect to all other causal factors, there is a second strategy to deal with Simpson's paradox – simply deny the phenomenon. I suppose this is not a crazy proposal because very clever people like Clark Glymour, Peter Spirtes, Judea Pearl and a host of followers insist on it. It is a major assumption in the currently fashionable Bayes-nets methods for causal discovery and causal inference. The chief argument offered in its support is one about mathematical spaces. Notice that in the Salmon example, only some specific arrangements of probabilities will make $P(Y/X) = \mathrm{P}(Y/-X)$. This is characteristic of Simpson's paradox situations. I can construct millions of them,

[13] Cartwright 1983

[14] Grunbaum 1993, p. 242

but always one has to get the numbers to balance out correctly. This is reflected in a theorem: if you put a Lebesgue measure over n-tuples of numbers, the set of n-tuples that gives rise to Simpson's paradox has measure zero. From this it is concluded that these situations never occur in nature. Not only is this a bad inference; the conclusion seems palpably untrue as well. I shall return briefly to this point in a moment.

There is a second well-known reason why a cause may not increase the probability of its effect: a particular cause may itself have different capacities with respect to the same effect and these capacities may balance out. This is especially likely to happen in systems we design ourselves, either consciously or by trial and error. One and the same cause can have opposing tendencies – both to enhance an effect and to retard it. In any case, where these tendencies just balance out, we will see no increase in the probability of the effect on the cause. Consider, for instance, a certain brand of non-drowsy decongestant: it does not put you to sleep. Nevertheless the patented chemical in it is a powerful soporific. The decongestant does not induce drowsiness because the chemical is always packaged with an equally potent stimulant.[15]

For cases like this where there is no change in the probability of the effect whether the cause is present or not, it may nevertheless be extremely useful to know the specific causal facts. Armed with the information about the Janus-faced nature of a cause, we may, for instance, be able to place some block in one of its pathways and thus be left with only the result of the tendency in the other direction.

Those who deny that Simpson's paradox situations ever occur generally also deny that cancellations of opposing tendencies will ever lead to equality of the conditional probabilities. The primary argument is the theorem as before: the numbers that afford exact cancellation form a set of measure zero. Again, this argument seems both invalid and unsound. Exact cancellations are often just what we try to achieve. For the decongestants, for example, they are supposed merely to decongest – they are supposed *neither* to put one to sleep nor to stimulate one. So exact cancellation matters.

Economic methodologist Kevin Hoover also takes this kind of cancellation to be common:

> Spirtes *et al.* (1993, p. 95) acknowledge the possibility that particular parameter values might result in [exact cancellation], but they dismiss their importance as having "measure zero". But this will not do for macroeconomics. It fails to account for the fact that in macroeconomic and other control contexts, the policymaker aims to set parameter values in just such a way to make this supposedly measure-zero situation occur. To the degree that policy is successful, such situations are common, not infinitely rare.[16]

It would also be a mistake to put too much emphasis on questions of whether

[15] Thanks to Lisa Lloyd for this example.
[16] Hoover 2001, p. 170

*exact* cancellations occur regularly or not for a reason central to the genuine practical considerations that underwrite Grunbaums's concerns about psychoanalytic claims. We are not primarily concerned with pure abstract philosophy – is increase of probability a necessary condition for causality or not? Rather we are interested in real methodologies that can be used to give reasonable confidence in claims that support often costly strategies for relief from very severe problems.

Our tools of statistical estimation are never, even in principle, good enough to settle a claim about exact equality of conditional probabilities. I think we have good empirical evidence that both near-Simpson's paradox situations and near-cancellations occur regularly and we know that our best methods will frequently estimate conditional probabilities as equal. If we adopt increase in probability as a necessary condition for a generic causal relation, we will give wrong verdicts about these. And this really matters because in both cases there may be large numbers of individuals for whom the cause has a profound effect.

This is my case against increase in probability as a necessary condition for generic level causation. Perhaps, however, Grunbaum might want to use it directly as a test for singular causation. Maybe not though. I myself have never been comfortable with talk of single-case probabilities – and I think I learned this suspicion from Grunbaum himself! They at any rate just push the problem back a step – how *for the single case* can we establish probability claims in a psychoanalytic setting?

Also I suspect we would frequently get the wrong verdicts. Consider one of those cases of opposing tendencies that cancel, where the cancellation in probabilities takes place not because different members of the population experience opposing outcomes but rather because the opposing tendencies balance out in each individual. Gerhard Hesslow's example of birth control pills and thrombosis is, I believe, supposed to work like this.[17] The pills cause thrombosis; they also prevent pregnancy, which itself is a cause of thrombosis. If the two tendencies cancel for an individual, say me, then I suppose we might find, for single case probabilities, $P$(N gets thrombosis/N takes pills) = $P$(N gets thrombosis/N does not take pills). Still we would want to know *both* causal facts because we may be able to find a way to block the deleterious pathway and leave me with only the beneficial effects (thereby changing the conditional probabilities).

The more usual way to use probabilities to treat singe cases is in a two-step process. We use the probabilities to establish generic causal claims, then – like Donald Davidson – we insist that all admissible singular claims fall under a generic law. So i) if we take increase in probability as a necessary condition for a generic causal truth, "$C$'s cause $E$'s" and ii) we take as a necessary condition for "this $C$ caused this $E$" that it be generically true that $C$'s cause $E$'s, then we would have as a result that increase in probability of $E$ on $C$ is a necessary condition for the related singular causal claim. I have already explained why I do not accept the first step. I am also suspicious of the second, whether or not we accept the first. I particularly worry about the second in medical and

[17] Hesslow 1976

psychiatric settings of the kind that Grunbaum is concerned with.

Let us begin anecdotally. I offer Emily a regime of diet to lose weight. She says, "That kind of crash diet never works". I reply, "It works for me". And I do so with good reason. I am fairly careful about these matters. I have tried this diet frequently. It always works. I know about the possibility of spontaneous weight loss and about the placebo effect, about long-term vs short-term outcomes *etc.* and I have evidence these are not a problem. Now there *may* be some description, $D$, of me that fixes *in a law-like way* the efficacy of this diet for people who satisfy $D$. In that case we would have a reference class, picked out in a non question-begging way, in which $P$(weight loss/diet) > $P$(weight loss/no diet). But there may not be such a description. There may simply be individual variation. To insist that there is always such a description is to let a big – and insecure – metaphysical assumption guide our methodology. That I think is a wrong thing to do.

It is however widely assumed. For instance consider Hersen and Barlow's text, still in use, *Single Case Experimental Designs.* Here they endorse this view, citing in turn the earlier text of M. Sidman on experimental data in psychology:

> Physics assumes that variability is imposed by error of measurement or other identifiable factors. Experimental efforts are then directed to discovering and eliminating as many sources of variability as possible [...] Sidman proposes that basic researches in psychology adopt this strategy. Rather than assuming that variability is intrinsic to the organism, one should make every effort to discover sources of behavioral variability among organisms [...][18]

Here of course we do not see the strong metaphysical position – variability is impossible, but merely a methodological injunction. But we need to be cautious even about methodological injunctions. We want to look for sources of variations because that knowledge could be powerful. But every hunt is costly, especially if we do not have good starting ideas of where to look. And in the case of human beings – who are not after all electrons – we cannot rule out the possibility of intrinsic variability. So we may be hunting for what in fact can never be found. In cases like this, we need to weigh the costs and the benefits, the probabilities of success and the probabilities of failure from following the advice.

The question is a live one, particularly now regarding pharmaceuticals. Some drugs may work very well for some people but for others may worsen the very condition they are supposed to treat. Ideally we would like a testable description for those in the two categories; and we would like to ensure that within these two categories, the treatment passes all the tests for generic-level causes. But such testable descriptions are often not available, and we must even leave open the possibility that they simply do not exist. The good news is that sometimes the bad effects set in gradually and in some known pattern so that patients can be monitored and taken off the treatment if it is proving harmful.

[18] Hersen and Barlow 1981, p. 35

One possible example that has recently received notoriety is Prozac, which is used to treat depressives. It seems, however, sometimes to induce akathesia, a kind of restlessness that can lead to suicide or even the killing of others. So far nobody knows how to predict who will fall in the category where Prozac increase the chance of suicide (or even if there is such a proper category). The controversy has arisen in part because many previous anti-depressants were recommended to be taken with tranquillisers; this was not so for Prozac – except in Germany; and now there is the suspicion that a number of people have killed themselves and others on account of taking the anti-depressant intended, *inter alia*, to prevent suicide.

There is a mire of legal, social, medical, moral and methodological problems here. It is commonplace that standard clinical trial procedures, even when the trial population is large and compliant, do not tell us when treatments have opposing effects. I have not found anything yet in Grunbaum's own writing that looks at the methodology in cases like this. One hope I have in raising these issues is that we can encourage him to turn to this tangle of problems to help sort them out, as he has so many other issues of importance to human welfare.

# References

Cartwright, Nancy 1983: *How the Laws of Physics Lie*, 1983, Oxford: Oxford University Press.

Cartwright, Nancy 1989: *Nature's Capacities and their Measurement*, Oxford: Oxford University Press.

Grunbaum, Adolf 1993: *Validation in the Clinical Theory of Psychoanalysis, A Study in the Philosophy of Psychoanalysis*, Madison, CT: International Universities Press.

Grunbaum, Adolf 1990: "'Meaning' Connections and Causal Connections in the Human Sciences", *Journal of the American Psychoanalytic Association* **38**, pp. 559-577.

Herson, Michael and David Barlow 1981: *Single-Case Experimental Designs*, Oxford: Pergamon Press.

Hesslow, Gerhard 1976: "Discussion: Two Notes on the Probabilistic Theory of Causality," *Philosophy of Science* **43:**2, pp. 290-2.

Hoover, Kevin 2001: *Causality in Macroeconomics*, Cambridge: Cambridge University Press.

Salmon, Wesley 1971: *Statistical Explanation and Statistical Relevance*, Pittsburgh: University of Pittsburgh Press.