Grünbaum's challenge to causal inference in psychoanalysis: 25 years on

Michael Lacewing

Adolf Grünbaum famously attacked what he called the 'foundations' of psychoanalysis. His claim was not that the psychoanalytic theory of the mind is untrue, but that the methodology by which psychoanalysts have constructed a central part of that theory, viz. the claim that neurosis is the result of repression, is unsound.[i] Psychoanalysis cannot be justified by the psychoanalytic method of theorizing from clinical data. After briefly examining the prospects of defending psychoanalysis through meeting of Grünbaum on his terms, I will object that Grünbaum's theory of causal inference is overly restrictive, and that Freud's reasoning corresponds legitimately to a different model, viz. inference to the best explanation. However, I will conclude that drawing on clinical data alone cannot secure inferences to the best explanation, and so I support Grünbaum's conclusion that the justification of the psychoanalytic model of the mind must rely in part on evidence outside clinical data.

1. Clinical data comprises the totality of patients' verbal and non-verbal behaviour in the clinical setting; once Freud developed the method of free association, this behaviour was (ideally) generated by the patient's free associations to whatever topic first entered their mind at the start of the session, including their free associated responses to the analyst's interpretations. According to Grünbaum (1984: 6-7)[ii], it was from this data, including evidence it provided of therapeutic success, that Freud inferred that neurosis had its origin in the repression of an experience or memory. The inference is reconstructed in §2 below.

In his discussion, Grünbaum focuses primarily on whether inferences based on free association can establish Freud's original (pre-1895) hypotheses about the etiology of neuroses. But there are at least three ways in which Freud's original hypotheses were distinct from his later theories of neurosis. First, he supposed the cause of neurosis to be repression of the memory or experience of an actual (non-psychical) event in the life of the subject; by 1897, Freud had come to argue that the cause of neurosis may itself be intra-psychical. For example, he had thought that actual sexual abuse was a necessary precondition of hysteria, but came to believe that a (childhood) fantasy of sexual abuse could play the necessary role. Second, Freud originally supposed that different specific neuroses (conversion hysteria, phobia, obsessional neurosis) had different specific causes (types of event). By the mid-1920s, however, he believed that all neuroses resulted from an inadequate resolution of the Oedipus complex, and gave up the attempt to explain why this rather than that neurosis was the result for any particular individual. The explanation is holistic rather than specific. Third, the original hypotheses all rested on an early theory of repression that was gradually refined into a generalized theory of psychological defence, within which 'repression' came to name a specific defence mechanism, one among many others. Neurosis, then, implicated psychological defence, but no longer implicated repression as opposed to projection, certain forms of identification, reaction formation, regression, or isolation.

As a result of these important changes in Freud's etiological account of neurosis, it seems strange to many that Grünbaum – while being perfectly aware of the changes – decided to critically discuss Freud's earliest theories. Grünbaum presented arguments directly aimed at hypotheses that psychoanalysts themselves had rejected by the 1920s as insupportable, both by the clinical data and by theoretical developments. In a sense, at least part of Grünbaum's critique was already sixty years out of date. And so many psychoanalysts and a few philosophers have responded to Grünbaum by arguing that Grünbaum misunderstands and misrepresents psychoanalysis. While I believe there is considerable truth in some of the charges laid against him, it is also true that Grünbaum himself is often misunderstood.

His view is this: The challenge he raises is whether psychoanalysis can validate its causal claims about the origins of neurosis by its own methods. According to Grünbaum, Freud based his first theories of how neurosis originated on clinical data alone, rather than, say, on epidemiological data or longitudinal studies, and many psychoanalysts continue to formulate and defend their causal hypotheses on the basis of clinical data. However, Freud's early theories had the greatest chance of being able to be validated, because the later changes all *make it much harder* to validate causal claims (264). As an example, while we might use evidence from memory to establish the occurrence of a particular childhood event, we cannot expect memory to furnish evidence of the occurrence of a childhood fantasy, all the more so if the fantasy were unconscious. Or again, the claim that a particular type of event is causally necessary for a particular type of neurosis is easier to establish (or refute) than the claim that it has some causal relevance. If it can be shown that Freud's most testable hypotheses cannot be validated by psychoanalytic methods, this is sufficient to undermine later developments as well (246).

Unfortunately, Grünbaum takes few pains to make this case. In part, this is because he takes the case to be straightforward – it is harder to establish the existence and causal relevance of purely psychological 'events' than of physical ones; it is harder to establish causal relevance than (stronger) causal necessity; it is harder to test an etiology that does not posit a specific or distinct (type of) cause; and so on. Methods of validating causal claims that do not suffice in the easier cases are not going to suffice in the more difficult cases, it would seem. Correcting Grünbaum's misunderstandings of psychoanalytic accounts of neurosis may provide no answer to his charge that the methodology on which such accounts rest is flawed.

2. Grünbaum holds that the appropriate standards for justifying causal inferences are the 'time-honored canons' of Bacon and Mill (1904). Of Mill's four canons (Agreement, Difference, Residues, Concomitant Variations), it is those of Agreement and Difference that are of most importance here. We employ the Canon of Agreement when, finding only one antecedent shared by all instances of some effect, we infer that to be a cause. We employ the Canon of Difference 'when we find that there is only one prior difference between a situation where the effect occurs and an otherwise similar situation where it does not, [and] infer that the antecedent that is only present in the case of the effect is a cause' (Lipton 2004: 18). So, according to these canons, to establish that X is the cause of Y, we must show that Xs make a difference to the occurrence or incidence of Ys *by comparing (classes of) cases*, either comparing the incidence of Ys in cases in which X occurs with ones in which it does not (Difference) and/or by comparing cases in which Y occurs to see whether X occurred or not (Agreement).

An example: to establish whether smoking causes lung cancer, we must compare the incidence of lung cancer in smokers and non-smokers; and/or we must examine cases of lung cancer to see whether the person smoked or not. If there is a higher probability that a smoker will have lung cancer than a non-smoker, we may infer that smoking causes lung cancer. Likewise if there is a higher probability that someone with lung cancer is a smoker than non-smoker.

It is important to note that this notion of 'cause' is simply that of a factor that *makes a difference*. There is no commitment that all causes operate according to laws. Mental states, such as beliefs, desires, wishes, phantasies, memories, and the like, make a difference to how we behave, though there may be no laws that govern the relations between such states and their effects. As such, there is no reason to suppose that psychological explanations are not either themselves causal explanations, or at least run in parallel to causal explanations.

On the face of it, inferring that the cause of neurosis is the repression of a memory on the basis of clinical evidence does not meet the standards of evidence of Mill's canons. For instance, while it may be possible, though unlikely, that clinical data could show that the only thing all neuroses have in common is repression, it certainly could not show that the difference between

neurosis and its absence is the occurrence of repression. On the assumptions that a) there are people who are not neurotic; and b) people who are not neurotic do not seek psychoanalytic help; clinical data tells us nothing about people who are not neurotic.

Anyone familiar with psychoanalytic theory will realize just how alien this approach is, with its presumption that people may fall into two classes – neurotic and non-neurotic. But this simply entails that if psychoanalysis is to legitimately draw causal inferences, it must defend a different methodology. I will argue in §4 that Grünbaum is mistaken in his belief that Millian canons are always needed to justify causal inference. Meanwhile, Freud needs a defence of his method of drawing causal inferences without extraclinical statistical studies, and Grünbaum claims to find two arguments in support.

The first, which Grünbaum named the 'Tally Argument', is this:

> (1) only the psychoanalytic method of interpretation and treatment can yield or mediate to the patient correct insight into the unconscious pathogens of his psychoneurosis, and (2) the analysand's correct insight into the etiology of his affliction and into the unconscious dynamics of his character is, in turn, *causally necessary* for the therapeutic conquest of his neurosis. (139-40)

If these two claims, which Grünbaum calls jointly the 'Necessary Condition Thesis' (NCT), are true, and if psychoanalysis is therapeutically successful, then this success validates the etiological hypotheses of psychoanalytic theory. The interpretations of the analyst must 'tally' with the real causes of the analysand's neurosis, or cure will not follow. Given that psychoanalysis provides therapeutic success, its methods are a means of discovering the causes of neurosis. Grünbaum intends this to apply equally to Freud's early claim that the cause of neurosis is a repressed memory/experience of some external event; or to later views that neurosis is caused by repressed phantasies or other intra-psychic structures.

Grünbaum takes this argument to be Freud's main response to the two challenges Grünbaum presents to psychoanalysis. The first is that we cannot know that clinical data, produced by free association, are not contaminated by suggestion. If data are contaminated, then inferences based on the data will, of course, be unable to justify those inferences. At the same time, therapeutic effects may be the result of suggestion, not genuine insight. Free association would therefore not be a method for discovering the true causes of neurosis, and NCT is false. However, I shall discuss suggestion no further here. My conclusions, therefore, will be provisional upon finding a satisfactory response to the question of suggestion in psychoanalytic psychotherapy.

Setting aside the issue of suggestion is not unfair to Grünbaum, as he states (128) that even if we knew clinical data were uncontaminated, psychoanalysis still could not validate its causal inferences, because its method of doing so falls prey to his second challenge, viz. that, if the Tally argument fails, the causal reasoning itself is flawed twice over.

First, Breuer and Freud defended the claim that repression caused neurosis on the grounds that lifting a repression was followed by the removal of a symptom. On this basis, they inferred that the symptom removal was caused by lifting the repression; and from there, that the original repression had caused the neurosis, which manifested in the symptom (among others). But unless NCT can be established, it cannot be shown that it was *lifting the repression* which led to the removal of the symptom, rather than some other event – such as suggestion – that was the cause. And so it cannot be inferred that repression caused neurosis in the first place.

Second, in developing his method of free association in lifting repressions, Freud came to believe that free association from neurotic symptoms led to the repressed memories that were

the cause of the neurosis (or again, free association from the manifest content of a dream to the wish that caused it, from a slip to its cause, and so on). The reasoning is thus: the patient free associates from a neurotic symptom; the psychoanalyst interprets; where the interpretation (together with the process of 'working through') lifts a repression, the symptom is removed; repression therefore causes neurotic symptoms; free association therefore leads from effect to cause. The claim that what comes 'after' in free association (the terminus point, the thought or wish that is interpreted from the material) is the cause of what comes 'before' (the starting point, and the material produced in free association) is therefore founded on the previous causal inference, from cure to the repression etiology of neurosis. But again, if NCT is false, then free association does not uniquely lead to cure through insight into the causes of neurosis, and so we could not conclude that free association is a method for discovering causes (of neurotic symptoms, dreams, slips).

Grünbaum argues that the Tally argument fails as at least the second premise of NCT is empirically unsupported – therapeutic success does not require correct etiological insight, as other therapies can be successful and there is even spontaneous remission. Again, Freud recognised this by the 1920s. But giving up the Tally argument means psychoanalysis needs a different defence of its practice of making causal hypotheses on the basis of clinical data. We shall examine Freud's second argument in §3.

Many critics of Grünbaum have argued that the Tally argument is not to be found in Freud. My concern is not Freudian exegesis, but the defensibility or otherwise of psychological causal hypotheses that are not established through methods of controlled experiment and/or statistical analysis. If the Tally argument is in Freud, but does not work, that does not avail psychoanalysis. If the Tally argument isn't in Freud, but is nonetheless true, it is a powerful defence of psychoanalytic theorizing.

One response to Grünbaum's challenges to causal inference in psychoanalysis, therefore, is to defend the Tally argument. Or rather, a weakened version of it, since there is no evidence to support that psychoanalytic therapy is causally necessary for 'cure'. But medical science does not work with necessary causes, only claims of causal relevance. There are two challenges to Tally argument: spontaneous remission and treatment by non-psychoanalytic therapies. To establish that psychoanalysis has a stronger therapeutic effect than absence of treatment and rival therapies would go some way to supporting the Tally argument.

Compared with spontaneous remission (no treatment), recent studies show that psychotherapy (in general, not psychoanalysis specifically) has a considerable positive effect, with 80% of those treated ending up better than no-treatment controls. This effect size is as large as many medications for physical complaints, and greater than almost all treatments in cardiology, geriatric medicine, asthma, flu vaccination and cataract surgery. However, when it comes to comparing rival psychological therapies, there is little evidence that psychoanalysis is superior. As not all psychotherapies employ free association, nor seek therapeutic success through insight into the etiology of neurosis, the evidence needed to establish NCT is missing.

But there are two reasons to think that this is more an absence of evidence that psychoanalysis has a greater therapeutic effect than evidence of absence of such an effect. First, the vast majority of outcome studies compared rival psychotherapies with *short-term* psychoanalytic psychotherapy (up to six months) rather than the traditional open-ended long-term psychoanalysis that normally lasts several years. But some recent studies (Leichsenring & Rabung (2008); Maat, de Jonghe, Schoevers & Dekker (2008)) indicate that the benefits of psychodynamic psychotherapy become much more pronounced after six months' treatment. Second, very few studies have measured what it is that psychoanalysis *now* would claim to be uniquely best at achieving. Outcome studies need something clear to measure, and it is less

expensive to measure it over a short period of time, but psychoanalysis attempts to go beyond overt 'symptom reduction', e.g. in improving the patient's ability to tolerate a wider range of emotional experiences and be more emotionally 'alive', to have a more satisfying sex life, to understand themselves and others in more nuanced ways, to live with greater freedom and flexibility. These therapeutic outcomes have rarely been measured. I believe the empirical work is yet to be done on establishing whether there is an appropriate interpretation of 'therapeutic success' according to which psychoanalysis uniquely delivers such success.

Suppose we discover evidence that long-term psychoanalytic psychotherapy has a greater therapeutic effect. This would not yet tell us *why* it has such effect. We would require further studies – of precisely what kind is difficult to envisage – to establish that it is insight into the analysand's psychic structures of defence rather than suggestion that is responsible for the therapeutic success. We cannot argue that the Tally argument succeeds, as the relevant evidence is simply missing, and outcome studies will need to undergo significant changes in design to take into account long-term psychoanalysis, the outcomes beyond overt symptom reduction that psychoanalysis seeks to promote, and the causal factors relevant to bringing about these effects.

3. Freud's second argument is an argument from convergence and consilience. 'Consilience', coined by William Whewell, is 'when an Induction, obtained from one class of facts, coincides with an Induction obtained from another different class. Thus Consilience is a test of the truth of the Theory in which it occurs.' (Whewell 1840: Vol. 2, p. 469) Or, a slight variation, consilience is when an induction based on one class of facts is able to provide an explanation of a different set of facts. Inductions based on each class of facts 'converge' on the same explanation or hypothesis.

While everyone agrees, I believe, that the argument is genuinely Freud's, there is disagreement over its success, its content, its scope, and its chronological placement. In his main commentary on it, Grünbaum locates it very late, taking it from Freud's (1937) paper 'Constructions in analysis'. Here Freud argues that the analyst's judgment that an interpretation is correct rests not directly on whether the analysand agrees or disagrees with it. It rests on other evidence as well, including non-verbal behaviour and what further free associations follow the interpretation. Interpretations relating to neurotic symptoms, to dreams, to slips, and to the transference all also need to cohere with one another, pointing to the same underlying structure of the neurosis. Grünbaum therefore takes the scope of the argument to relate just to these data, and its content to concern clinical interpretations of each analysand individually (in the first instance).

Grünbaum objects that the argument fails to support psychoanalytic theorizing. First, the data are not independent if they are all equally contaminated by suggestion, so their consilience with one another provides no support for the truth of the interpretations made (277). Consilience only has probative weight when the consilient data are genuinely independent. Second, even genuine consilience would not establish causal hypotheses. The consilient data is all produced by free association; but if NCT cannot be established, we have no reason to think that free association leads to the discovery of the causes of neurosis. Only the use Millian canons can establish causal inferences.

One response, defended in detail by David Sachs (1989), is that Freud's argument is much more wide-ranging. Freud defended his causal hypotheses as of a piece with his general structural account of the mind, which both draws upon and provides successful explanations of many phenomena not primarily encountered in the clinical setting, such as mythology, jokes, religious belief, moral obligation, and biography, among others. As such, not all the data depend upon free association and the data are independent of each other. However, Grünbaum responds that Freud's interpretations of non-clinical phenomena themselves rest on the clinical phenomena. Therefore, the consilience between clinical data and psychoanalytic explanations of non-

psychoanalytic phenomena is again spurious, because the latter are not independent of the former.

We could expand the range of data yet further in search of independence. For example, Neil Manson has persuasively argued that Freud's early theorizing was informed by his 'economic model' of psychic energy,[iii] and we may argue that the development of this model rested on, at least in part, neurological and biological data that *was* independent of the clinical data. But the economic model is not a good basis for defending Freud, as there is little evidence that it is correct (Holt 1954?). We would do much better mounting the argument drawing on different sources, such as neuroscience and child development theory (Weston).

This may help establish the truth of psychoanalytic claims, but helps not at all against Grünbaum's argument. Grünbaum is not interested in demonstrating that Freud was *wrong*, only that his conclusions have not been well-established on the basis of clinical data alone. Any appeal to consilience with extraclinical data grants Grünbaum's main contention, viz. that clinical data *on its own* is insufficient to justify causal inferences. Psychoanalysis must rest on discoveries outside psychoanalysis; it cannot be justified from within.

To recap: The first line of defence against Grünbaum's attack on the 'foundations' of psychoanalysis was to rehabilitate the Tally Argument. We concluded that this line of argument is currently unavailable. The second line of defence was to appeal to the consilience of clinical data. To this, Grünbaum objects that we must first demonstrate that free association is a method that reliably discovers causes. It is not, he argues, as causal inference requires a different form of reasoning. In what follows, I argue that, quite independently of any special considerations relating to psychoanalysis, Grünbaum holds a mistaken theory of causal inference. I then discuss the claim that on a corrected theory of causal inference, psychoanalytic methodology is defensible.

4. As a first step in examining the issue of causal inference, we should clarify Grünbaum's objection that even genuine consilience in clinical data would not secure causal inferences. The difficulty here is with establishing causal connections between the material generated by free association. The causal connections are inferred, Grünbaum comments, on the basis of what he terms 'thematic affinity', or what we might otherwise term 'connections of Intentional content'. Such connections can be observed in a stream of practical reasoning: I desire *a drink of water*; I believe that *water is available to drink in the kitchen*; I form the intention *to go to the kitchen*. Breuer's patient Anna O. suffered from hydrophobia – even as she was thirsty, she could not bring herself to drink water. Under hypnosis, she recalled seeing the dog of her English lady companion (towards whom she felt suppressed anger) drink from her companion's water glass, at which she felt strong disgust. Upon awaking from hypnosis, she asked for and drank a glass of water, her disgust now gone. Breuer inferred that the repression of her memory of the incident, and the feelings attached to it, caused her hydrophobia. The inference was based on the Intentional content of the relevant states (the perception, the anger, the disgust, the memory, the hydrophobia), states recalled under hypnosis, which Freud later replaced by the method of free association.

Grünbaum argues that causal inferences from an action or mental state to another mental state cannot be justified on the grounds of 'thematic affinity' alone. So, for instance, we cannot infer from your going to the kitchen to get a drink of water alone that the cause of your action is a desire for a drink of water. Again, we cannot infer that what comes after in free association is caused by what comes before (or vice-versa) just on the grounds that the two are linked by (chains of) thematic affinity. In making the argument, he repeats his contention that only Millian canons can justify causal inferences. To repeat, this objection to free association affects the argument from consilience. If the data involved is all clinical data, then all the causal hypotheses

involved will be based on the thematic affinities of free association. As this ground of causal inference is illegitimate, consilience between such inferences will provide no additional support.

Let us contrast Grünbaum's claim that the appropriate methods for testing causal hypotheses (always, and therefore also in psychology) derive from the eliminative inductivist methods of Bacon and Mill with an examination of *inference to the best explanation*. We will then discuss the claim that psychoanalytic inferences on the basis of thematic affinity can be justified as a form of inference to the best explanation.

The most extended development and defence of inference to the best explanation is Peter Lipton's (2004) book. Lipton argues that inference to the best explanation is, in most cases, a better account of scientific causal reasoning than rival accounts (and that it also cracks many of the standard philosophical chestnuts that perplex its rivals). I shan't repeat his defence here, but it is worth briefly contrasting with Mill's canons.

Summarily, then, inference to the best explanation uses explanatory considerations to support the inference. We only consider causal hypotheses (consistent with the evidence) that, if true, provide a plausible explanation of the effect; and we select from among these the causal inference we take to provide the best explanation of the effect. So in drawing causal inferences, we rely on such considerations as whether the hypothesis, if true, would provide a unifying account of phenomena, taking into account its scope, simplicity, and consilience with other evidence; whether it fits well or poorly with existing background knowledge; and, in many cases, whether a causal mechanism might be provided to connect the hypothesized cause with its effect.

Lipton's discussion provides at least three reasons to prefer an account of causal inference as inference to the best explanation over Mill's canons. First, Mill's four methods (Agreement, Difference, Residues, Concomitant Variations) are 'agreed on all sides to be an incomplete account of inductive inference' (Lipton 2004: 126). Even Mill allows that they must be supplemented by the Method of Hypothesis (Mill 1904: III.XIV.4-5). The problem is illustrated by Hempel's (1966: 15) remark that coming up with good hypotheses is ultimately a matter of 'happy guesses'. Grünbaum remarks (189) that it is entirely possible that while Freud's methods were epistemically flawed, his brilliant clinical mind may yet have landed on the truth – which suggests much the same view as Hempel's, of hypotheses being generated independently of evidence. Inference to the best explanation, by contrast, provides a more principled account of hypothesis creation, viz. that it is guided by explanatory concerns, and connects the context of discovery to the context of justification, both being governed by concern for securing the best explanation.

Second, Mill accepts that the all-important Method of Difference is idealized: there is never *just one* difference between cases. We may ignore differences we know to be irrelevant, says Mill. But first, experimentation may still introduce more than one difference, so further principles of selection are still needed; and second, we need to account for the origin and the role of our background knowledge regarding which differences to ignore (Lipton 2004: 128). Inference to the best explanation provides a good account of the principles of selection and background knowledge in judgments of relevance. We look for a difference that, if it were causally relevant, would provide the best explanation of the evidence we have.

Third, a second limitation is that the Method of Difference is unable to account for inference to *unobserved* or *unobservable* causes (Lipton 2004: 127). It only licenses the inference that of some effect, a particular observed antecedent is a cause. It does not say anything about the *discovery* or *inference* of an antecedent difference in the cases being investigated. Inference to the best explanation, by contrast, allows one to infer that there is some unobserved causal difference

between the cases, since precisely this would best explain why the cases turn out differently (as observed). The Method of Difference compares the known histories of the two cases to see if there is a difference. Inference to the best explanation raises the additional question of whether there being a difference would best explained the contrastive evidence. If so, we infer that the difference exists and is a cause.

As an application of this, inference to the best explanation in general deals well with 'self-evidencing' explanations, as in explaining why there are footprints in the snow by inferring that someone passed this way (Lipton 2004: 56). The best evidence that someone passed this way are the footprints themselves. This introduces a circularity: what explains the footprints? someone passing; how do we know someone passed? the footprints. But the circularity is benign, and inference to the best explanation accounts for why this is. The issue of self-evidencing explanations to unobserved or unobservable causes is clearly *particularly* relevant in the case of psychology, where the causal inferences we are frequently drawing are to unobservable mental states, and where we take some of the best evidence for the existence of such mental states to be the evidence from which we make the causal inference itself.

A fourth improvement Lipton does not comment on. Mill makes it clear that his Methods (seek to) establish causal laws, and the method of comparison clearly indicates they are most appropriate for type-type relations between cause and effect. This makes them problematic as an account of singular causal inference. We clearly employ singular causal explanation ('why are you late?', 'why doesn't the car start?', 'why does she believe in God?'), and we do so in ignorance of the laws, if any, that may exist governing the causal relations between cause and effect in this instance. Inference to the best explanation, however, is flexible enough to cope with singular causal inference; we draw the inference on the basis that if true, provides the best (singular causal) explanation. We can seek and provide an answer to 'why did P happen?' without necessarily talking about all instances of P, in whatever varied circumstances Ps occur, or comparing all instances of P to all instances in which P does not occur. If, in providing a singular causal explanation, we draw upon non-strict generalizations or causal schema in formulating the explanation, this only reinforces the claim that inference to the best explanation is the better account, since it provides explicitly for background knowledge.

5. I have given some initial reasons to suspect that inference to the best explanation may be a better account of causal inference than Mill's eliminative inductivism. We can use these reasons to argue more generally that even when eliminative inductivist methods are used, they usually form part of an inference to the best explanation, as they draw upon background knowledge and other explanatory concerns at least in the generation of hypotheses if not elsewhere. Lipton (2004: 121) grants that the conclusion that causal inference is *always* best understood as inference to the best explanation is too strong. But we may argue that at least many causal inferences in psychology are best understood in this way. First, causal inference in psychology is often inference to an unobserved or unobservable cause. Second, we often want to employ singular causal inference in psychology. The ability of inference to the best explanation to generate singular causal explanations without passing through type-type connections is important, as the difficulty of establishing laws in psychology is well-known.

Freud's argument from convergence and consilience is best interpreted as inference to the best explanation. Grünbaum argues that such an argument cannot support causal inference, on the grounds that this requires Mill's methods. But if, in general, inference to the best explanation can support causal inference, then Freud's argument may not be *methodologically* flawed and Grünbaum's argument, as it is formulated, must be rejected.

Furthermore, as Hopkins and others have noted, Freud's argument proceeds by inferring causal hypotheses on the basis of thematic affinity, and this form of reasoning has an important

precedent, viz. how we seek to understand people in everyday life. Our primary form of access to the motives of other people, and therefore to the causes of their behaviour, is that behaviour itself, including here what they say. There is a correspondence in content – thematic affinity – between behaviour and motive, between getting a drink of water and wanting to get a drink of water, between being repulsed by a drink of water and disgust at the thought of drinking water, and so on. To interpret the 'thematic content' of the behaviour is to place it in relation to a mental cause with just that content.

As Freud's argument notes, inferences of this kind are constantly open to revision, in light of further evidence: 'we constantly integrate the explanation we are inclined to give for on action with that we are inclined to give for others' (Hopkins 1988: 39). The more behaviour we seek to explain, the more complex our account of its mental causes. As we find the explanation for one aspect of behaviour matches or extends to other behaviours, so we have consilience. This process is driven, in large part, by thematic affinity. We are seeking an account of a causal system that best explains behaviour – we are engaged in inference to the best explanation on the basis of thematic affinity. The thematic affinities involved in psychoanalytic explanations are frequently different from those with which commonsense explanation works, as they work with chains of association rather than the more logical implications between beliefs and desires. But, Hopkins argues, the methodology is substantially the same.

As many causal inferences are drawn on the basis of clinical data, of varying levels of complexity, it is worth reminding ourselves of Grünbaum's target here, viz. psychoanalytic theories about the original causes of neurosis. These are high-level, general claims at some distance from the clinical data. To defend an inference to such a claim as the best explanation for neurosis will take a great deal more evidence and argument than, say, an interpretation of an analysand's current unconscious preoccupations or the themes of their psychic conflict. Defending the latter as the best explanation of the clinical data may succeed while a defence of the former may fail. It is an unfortunate mark of the debate that both sides have dealt with psychoanalytic causal inferences as though they were all of one form. Inference to the best explanation provides clear guidance on differentiating between causal claims, depending on explanatory power and assumptions.

The methodology of inference to the best explanation does not license drawing an explanation if there is insufficient reason to think that explanation is the 'best' or if no explanation is found that is 'good enough' (Lipton 2004: 63). Hopkins' restatement of Freud's argument from consilience faces the following objection, a restatement, in turn, of Grünbaum's original challenge:[iv] causal inferences regarding the origin of neurosis drawn on the basis of clinical data alone fail to meet the standards necessary for qualifying as the best explanation; for this, extraclinical data will be needed. In other words, even using inference to the best explanation, causal inferences in psychoanalysis may be methodologically flawed if they rely on clinical data alone. I shall conclude by arguing for this claim.

The most striking point in support of the objection is that for an inference to be to the *best* explanation, we need knowledge of other potential explanations. There are at least two 'environments' in which psychoanalytic explanations compete. It is noteworthy that very rarely are *either* of them referred to when causal inferences are being drawn.

The first, 'near' environment is the range of alternative psychoanalytic theories. In extending his critique from Freud's early theories of neurosis to the more recent claims of object relations theory and Kohut's self-psychology, Grünbaum asks (247) whether any theory can demonstrate that it is evidentially better supported than the others. The methodological strictures of inference to the best explanation require us to show that it is not just 'an' explanation, but the 'best' explanation. Yet it is a familiar caricature, which may yet have some truth in reality, that each

school of psychoanalysis draws inferences from similar clinical material in accordance with its own theories. What is necessary, but absent, is an account that demonstrates one set of inferences is superior to the others.

One response to this, from Rosenblatt (1989: 90-91), is that the differences are minor by comparison with what they agree upon, viz. that neurosis is causally related to unconscious defensive mentation. At the point at which Grünbaum discusses alternative psychoanalytic theories, he refocuses his critique, from the more general claim that repression causes neurosis to the more specific claim of Freud's that the repressed material is sexual in nature. Inference to the best explanation can allow that the more general explanation re. repression could be well-founded while the more specific explanations (Freudian, Kohutian, object relations) are not.

The second, more 'distant' environment in which psychoanalytic theories of neurosis compete comprises the range of non-psychoanalytic theories concerned with or impacting upon the data from which psychoanalytic inferences are drawn (e.g. alternative theories of dreaming and neurosis). Once again, to assess *any* inference as the best explanation, we need knowledge of the alternatives, and many of these may be generated outside the consulting room.

An objection may be raised that many of the phenomena that psychoanalysis seeks to explain simply do not have satisfactory alternative explanations, and that unifying power of psychoanalytic explanation continues to provide reason to accept it. However, we can only know this to be true, if it is true, *if we know what these alternatives are, and what the evidence is that supports them*. Both factors change constantly – what was the best explanation may cease to be so, either because new evidence is discovered that undermines the existing best explanation (e.g. through being recalcitrant) or because a new, more powerful explanation is generated. The 'environment of competition' can change, and has, of course, changed dramatically since Freud's initial proposals regarding the origin of neurosis 120 years ago (Fonagy).

These first points relate to knowledge of alternative explanations. A second set of considerations relates to the place of background knowledge in inference to the best explanation. To have confidence that a particular psychoanalytic causal inference is correct, we need to ensure that it does not directly conflict with well-established data elsewhere. It may that a causal hypothesis has implications, e.g. relating to childhood development or mental processes, which are undermined by non-clinical data. On the other hand, non-clinical investigations may provide corroboration for an explanation first arrived at from clinical data, either in its existing form or in a modified form. This interaction between clinical and non-clinical data refines the explanation that is finally justified as 'best'.

It may be, of course, that many of those psychoanalysts responsible for driving forward the development of psychoanalytic theory have been informed in their inferences in precisely the ways discussed, drawing on their knowledge of non-psychoanalytic psychiatry, child development theory, neuroscience, and other arenas of investigation. This does not undermine the objection, however. Psychoanalytic theories of neurosis have not been justified or developed on the basis of clinical data alone, but draw appropriately on a wider body of evidence.

(This relation of psychoanalytic theory to non-clinical data should not, of course, be seen as a one-way street. Other psychological explanations and theories are just as beholden to well-established psychoanalytic clinical data. There are famous difficulties involved in making this data available in a form that other theories may make use of, but the principles remain the same.)

6. To conclude: focusing his discussion on Freud's early hypotheses about the origin of neurosis, Grünbaum argues that Freud attempted to establish causal inferences on the basis of clinical data alone. Two challenges lead to the conclusion that this is a methodological flaw. The first,

which we have not discussed, is that we cannot establish that the data are uncontaminated by suggestion. The second is that Freud's causal reasoning is flawed. Grünbaum identifies two arguments in Freud to answer these challenges: the Tally argument and the argument from convergence and consilience. I argued that we cannot use the Tally argument due to lack of evidence. I noted that Grünbaum's objection that consilience cannot establish causal inferences rests on the view that causal inferences can only be established using Mill's Canons, but argued that this theory of causal inference is surpassed by that of inference to the best explanation. Therefore, if Freud's argument from consilience is a form of inference to the best explanation, it is not methodologically flawed in the way Grünbaum charges. Furthermore, arguments from Hopkins and others defend the claim that thematic affinity is a means by which causal inferences commonly are, and justifiably, drawn. However, this response to Grünbaum faces a distinct challenge, viz. that while inferences relating to an analysand's current unconscious mental states may well be justified, inferences regarding something as general as the origins of neurosis cannot be established as part of the best explanation on the basis of clinical data alone. For an inference to be to the *best* explanation, we need knowledge of other potential explanations and an understanding of how the inference fits with background knowledge. Both alternative explanations and background knowledge may be informed by non-clinical data; and therefore, so must psychoanalytic hypothesizing about general causal explanations of neurosis.

References

Freud, S. (1937) 'Constructions in analysis', SE

Grünbaum, A. (1984) *The Foundations of Psychoanalysis* (Berkeley: University of California Press)

Hempel, C. (1966) *The Philosophy of Natural Science* (Englewood Cliffs: Prentice-Hall)

Hopkins, J. (1988), 'Epistemology and Depth Psychology: Critical Notes on *The Foundations of Psychoanalysis*', in Clarke & Wright, *Mind, Psychoanalysis and Science* (Oxford: Blackwell), 33-60

Leichsenring, F. & Rabung, S. (2008), 'Effectiveness of long-term psychodynamic psychotherapy: a meta-analysis', *Journal of the American Medical Association* (300: 13): 1551-65

Lipton, P. (2004) *Inference to the Best Explanation* (2nd ed.) (London: Routledge)

Maat, de Jonghe, Schoevers & Dekker (2008), 'The effectiveness of long-term psychoanalytic therapy: a systematic review of empirical studies', *Harvard Review of Psychiatry* (17): 1-23

Manson, N. (2003) 'Freud's own blend', *Proceedings of the Aristotelian Society*

Mill, J. S. (1904) *A System of Logic* (8th ed.) (London: Longmans, Green and Co.)

Rosenblatt, A. (1989) 'Reinspecting the Foundations of Psychoanalysis: A Rejoinder to Adolph Grünbaum', *Psychoanalysis and Contemporary Thought* (12), 73-96

Sachs,, D. (1989) 'In Fairness to Freud: A Critical Notice of the Foundations of Psychoanalysis', *The Philosophical Review* (98), 349-78

Whewell, W. (1840) *The Philosophy of the Inductive Sciences*

[i] The claim stems from Freud's early theories. It is more accurately expressed now as 'neurosis is the result of chronic psychic conflict and the failure of psychic defence.

[ii] Page references are to Grünbaum 1984 unless otherwise indicated.

[iii] The economic model is a factor Grünbaum explicitly does not deal with, for two understandable reasons – Freud occasionally denies its centrality to psychoanalysis (1984: 5) (while elsewhere insisting on it!), and many psychoanalysts today reject the economic model (see Grünbaum 1984: 100-1).

[iv] Grünbaum has objected to Hopkins' account on other grounds as well, which I discuss elsewhere.