Retrospective Versus Prospective Testing of Aetiological Hypotheses in Freudian Theory

I. Introduction

The repression-aetiology of neurotic disorders is the cornerstone of the psychoanalytic theory of unconscious motivations. Repressions are held to be not only the pathogens of the psychoneuroses but also the motives of dream construction, and the causes of various sorts of bungled actions ("parapraxes") in which the subject is normally successful (e.g., slips of the tongue or pen). Thus even in the case of "normal" people, Freud saw manifest dream content and various sorts of "slips" as the telltale symptoms of (temporary) minineuroses, engendered by repressions.

Having modified Breuer's cathartic method of investigation and therapy, Freud arrived at the purported sexual aetiologies of the psychoneuroses, as well as at the supposed causes of dreams and parapraxes, by lifting presumed repressions via the patient's allegedly "free" associations. At the same time, excavation of the pertinent repressed ideation was to remove the pathogens of the patient's afflictions. Hence scientifically, Freud deemed the psychoanalytic method of investigation to be both heuristic and probative, over and above being a method of therapy.

Yet he and his disciples have not been alone in maintaining that the clinical setting of the psychoanalytic treatment sessions provides cogent epistemic avenues for testing his aetiologies retrospectively. The scrutiny of this thesis has highly instructive import for adjudicating the general polemic waged by critics against the clinical validation that Freudians have claimed for their theory. But the examination of the purported cogent testability likewise helps us assess the merits of prospective versus retrospective testing of aetiologic hypotheses in psychiatry. Hence I shall focus this essay on the scrutiny of the alleged clinical testability of the Freudian aetiologies.
II. Statement of the Controversy as to Clinical Testability

Hans Eysenck (1963) has maintained that "we can no more test Freudian hypotheses 'on the couch' [used by the patient during psychoanalytic treatment] than we can adjudicate between the rival hypotheses of Newton and Einstein by going to sleep under an apple tree." (p. 220) And, in Eysenck's view, only suitably designed experimental studies can perform the probative role of tests, while the clinical data from the couch may be heuristically fruitful by suggesting hypotheses. As against this denial of clinical testability, Clark Glymour (1974) has argued that "the theory Sigmund Freud developed at the turn of the century was strong enough to be tested [cogently] on the couch." (p. 304) And Glymour proposes to illuminate Eysenck's disparagement of clinical data but then to discount it by the following dialectical give-and-take:

It stems in part, I think, from what are genuine drawbacks to clinical testing; for example, the problem of ensuring that a patient's responses are not simply the result of suggestion or the feeling, not without foundation, that the "basic data" obtained from clinical sessions—the patient's introspective reports of his own feelings, reports of dreams, memories of childhood and adolescence—are less reliable than we should like. But neither of these considerations seems sufficient to reject the clinical method generally, although they may of course be sufficient to warrant us in rejecting particular clinical results. Clinicians can hopefully be trained so as not to elicit by suggestion the expected responses from their patients; patients' reports can sometimes be checked independently, as in the case of memories, and even when they cannot be so checked there is no good reason to distrust them generally. But I think condemnations like Eysenck's derive from a belief about clinical testing which goes considerably beyond either of these points: the belief that clinical sessions, even cleansed of suggestibility and of doubts about the reliability of patients' reports, can involve no rational strategy for testing theories.

I think that Eysenck's claim is wrong. I think there is a rational strategy for testing important parts of psychoanalysis, a strategy that relies almost exclusively on clinical evidence; moreover, I think this strategy is immanent in at least one of Freud's case studies, that of the Rat Man. Indeed, I want to make a much bolder claim. The strategy involved in the Rat Man case is essentially the same as a strategy very frequently used in testing physical theories. Further, this strategy, while simple enough, is more powerful than the hypothetico-deductive-falsification strategy described for us by so many philosophers of science. (p. 287-88)
Despite this epistemological tribute to Freud's couch, Glymour issues a caveat:

I am certainly not claiming that there is good clinical evidence for Freud's theory; I am claiming that if one wants to test psychoanalysis, there is a reasonable strategy for doing so which can be, and to some degree has been, effected through clinical sessions. (p. 288)

Most recently, Glymour (1980) told us more explicitly why we should countenance the rationale that animated Freud's clinical investigation of psychoanalytic hypotheses during the treatment of his Rat Man patient Paul Lorenz. Glymour points to at least three important specific episodes in the history of physical science in which he discerns just the logical pincer-and-bootstrap strategy of piecemeal testing that he also teased out from Freud's analysis of Paul Lorenz. Thus he says, "unlikely as it may sound . . . the major argument of the Rat Man case is not so very different from the major argument of Book III of Newton's Principia." (Glymour 1980, p. 265) And he stresses that this argument employs a logical pincer strategy of more or less piecemeal testing within an overall theory, instead of the completely global theory-appraisal of the hypothetico-deductive method, which altogether abjures any attempt to rate different components of the theory individually as to their merits in the face of the evidence.

Besides commending Freud's clinical study of the Rat Man for its rationale, Glymour likewise attributes a fair degree of scientific rigor to a few of Freud's other investigations. But he couples these particular appreciative judgments with a largely uncomplimentary overall evaluation, after deploring the very uneven logical quality of the Freudian corpus. Indeed, Glymour thinks he is being "gentle" (p. 265) when he deems Freud's 1909 case study of Little Hans to be "appalling." And he finds that "on the whole Freud's arguments for psychoanalytic theory are dreadful," marred by the frequent—though by no means universal—substitution of "rhetorical flourish" for real argument, and of a "superabundance of explanation" for cogent evidence. (p. 264) Yet clearly these quite fundamental dissatisfaction with Freud's all too frequent lapses do not militate against Glymour's espousal of the clinical testability of such central parts of psychoanalytic theory as the specific aetiologies of the psychoneuroses, at least in the aetiological versions that Freud enunciated before 1909.
Just this championship of the *probative* value of data from the analytic treatment sessions is philosophical music to the ears of those who echo Freud's own emphatic claim that the bulk of psychoanalytic theory is well founded empirically. For, as Ernest Jones reminded everyone in his Editorial Preface to Freud's *Collected Papers*, the clinical findings are "the real basis of Psycho-analysis. All of Professor Freud's other works and theories are essentially founded on the clinical investigations." (Jones 1959, vol. 1, p. 3) Thus most advocates of this theoretical corpus regard the analyst's many observations of the patient's interactions with him in the treatment sessions as the source of findings that are simply *peerless* not only heuristically but also probatively. We are told that during a typical analysis, which lasts for some years, the analyst accumulates a vast number of variegated data from each patient that furnish evidence relevant to Freud's theory of personality no less than to the dynamics and outcome of his therapy. The so-called psychoanalytic interview sessions are claimed to yield genuinely probative data because of the alleged real-life nature of the rich relationship between the analyst and the analysand. Even an analyst who just declared it to be high time that Freudians "move from overreliance on our hypothetical discoveries to a much needed validation of our basic theoretical and clinical concepts" (Kaplan 1981, p. 23) characterizes "the naturalistic observations within the psychoanalytic treatment situation" as "the major scientific method of psychoanalysis" (p. 18). Hence the clinical setting or "psychoanalytic situation" is purported to be the arena of experiments *in situ*, in marked contrast to the contrived environment of the psychological laboratory with its superficial, transitory interaction between the experimental psychologist and his subject. Thus the analysts A. M. Cooper and R. Michels (1978) tell us that "increasingly this [psychoanalytic] inquiry has recognized the analytic situation itself as paradigmatic for all human interactions." (p. 276) Indeed, the psychoanalytic method is said to be uniquely suited to first eliciting some of the important manifestations of the unconscious processes to which Freud's *depth* psychology pertains.

This superior *investigative value* of the analyst's clinical techniques is thus held to make the psychoanalytic interview at once the prime testing ground and the heuristic inspiration for Freud's theory of personality as well as for his therapy. Some leading orthodox analytic theoreticians have been concerned to *exclude* the so-called metapsychology of Freud's psychic energy model, and a fortiori its erstwhile neurobiological trappings from the avowed purview of clinical validation. Therefore it is to be
understood that the term "psychoanalytic theory of personality" is here construed to exclude the metapsychology of psychic energy with its cathexes and anticathexes. In any case, most analysts have traditionally been quite sceptical, if not outright hostile, toward attempts to test Freudian theory experimentally outside the psychoanalytic interview.

Just such an assessment was enunciated again quite recently by Lester Luborsky and Donald Spence (1978). They do issue the sobering caveat that "psychoanalysts, like other psychotherapists, literally do not know how they achieve their results." (p. 360) But they couple this disclaimer with the tribute that analysts "possess a unique store of clinical wisdom." Moreover, Luborsky and Spence emphasize that "far more is known now [in psychoanalysis] through clinical wisdom than is known through quantitative [i.e., controlled] objective studies." (p. 350, italics in original)

In short, they claim that—in this area—clinical confirmation is presently superior to experimentally obtained validation. And they deem findings from the psychoanalytic session to have such epistemic superiority not only therapeutically but also in the validation of Freud's general theory of unconscious motivations. (pp. 356-57) Similarly for the purported validation of Heinz Kohut's currently influential variant of psychoanalysis, which supplants Freud's Oedipal conflict by the child's pre-Oedipal narcissistic struggle for a cohesive self as the major determinant of adult personality structure. (Ornstein 1978)

Despite their strong differences, both of the parties to the above dispute about the probative value of clinical data for the empirical appraisal of psychoanalytic theory do agree that at least part of the Freudian corpus is indeed cogently testable by empirical findings of some sort: The Freudians have the support of Glymour, for example, in contending that observations made within the confines of the treatment setting do or can afford sound testability, while such anti-Freudian protagonists as Eysenck make the contrary claim that well-conceived tests are possible, at least in principle, but only in the controlled environment of the laboratory or in other extra-clinical contexts. And clearly the assumption of empirical testability shared by the disputants is likewise affirmed by someone who maintains that both clinical and extra-clinical findings are suitable, at least in principle, for testing psychoanalysis.

Yet this shared assumption of testability has again been denied simpliciter by Popper as recently as in his 1974 reply to his critics (Popper 1974, vol. 2, pp. 984-985). There he reiterates his earlier claim that Freud's
theory, as well as Adler's, are "simply non-testable, irrefutable. There was no conceivable human behaviour which would contradict them." (Popper 1962, p. 37) It is then a mere corollary of this thesis of nontestability that clinical data, in particular, likewise cannot serve as a basis for genuine empirical tests.

III. Can Popper's Indictment of Freudian Theory Be Sustained?

In earlier publications, I have argued that neither the Freudian theory of personality nor the therapeutic tenets of psychoanalysis is untestable in Popper's sense. (Grunbaum 1976, 1977, 1979a) Furthermore, there I contended in detail that Popper's portrayal of psychoanalysis as a theory that is entitled to claim good inductivist credentials is predicated on a caricature of the requirements for theory validation laid down by such arch inductivists as Bacon and Mill. Thus I pointed out that Freud's theory is replete with causal hypotheses and that the evidential conditions that must be met to furnish genuine inductive support for such hypotheses are very demanding. But I emphasized that precisely these exacting inductivist conditions were pathetically unfulfilled by those Freudians who claimed ubiquitous confirmation of the psychoanalytic corpus, to Popper's fully justified consternation.

In Grünbaum (1984, chap. 1, sec B), the reader will find my most recent detailed critique of Popper's charge of unfalsifiability against all of psychoanalytic theory. Hence it will suffice here to give mere illustrations of the textually untutored and logically slipshod character of this wholesale charge.

Even a casual perusal of the mere titles of Freud's papers and lectures in the Standard Edition yields two examples of falsifiability. And the second is a case of acknowledged falsification to boot. The first is the paper "A Case of Paranoia Running Counter to the Psychoanalytic Theory of the Disease" (S.E. 1915, 14:263-272); the second is the Lecture "Revision of the Theory of Dreams" (S.E. 1933, 22:7-30, esp. pp. 28-30). Let us consider the first.

The "psychoanalytic theory of paranoia," which is at issue in the paper, is the hypothesis that repressed homosexual love is causally necessary for one to be afflicted by paranoid delusions. (S.E. 1915, 14:265-66). The patient was a young woman who had sought out a lawyer for protection from the molestation of a man with whom she had been having an affair. The lawyer suspected paranoia when she charged that her lover had gotten unseen
witnesses to photograph them while making love, and that he was now in a position to use the photographs to disgrace her publicly and compel her to resign her job. Moreover, letters from her lover that she had turned over to the lawyer deplored that their beautiful and tender relationship was being destroyed by her unfortunate morbid idea. Nonetheless, aware that truth is sometimes stranger than fiction, the lawyer asked Freud for his psychiatric judgment as to whether the young woman was actually paranoid.

The lover’s letters made “a very favorable impression” on Freud, thereby lending some credence to the delusional character of the young woman’s complaints. But, assuming that she was indeed paranoid, Freud’s initial session with her led to a theoretically disconcerting conclusion: “The girl seemed to be defending herself against love for a man by directly transforming the lover into a persecutor: there was no sign of the influence of a woman, no trace of a struggle against a homosexual attachment.” (S.E. 1915, 14:265) If she was indeed delusional, then this seeming total absence of repressed homosexuality “emphatically contradicted” Freud’s prior hypothesis of a homosexual aetiology for paranoia. Thus, he reasoned: “Either the theory must be given up or else, in view of this departure from our [theoretical] expectations, we must side with the lawyer and assume that this was no paranoiac combination but an actual experience which had been correctly interpreted.” (S.E. 1915, 14:266) And furthermore: “In these circumstances the simplest thing would have been to abandon the theory that the delusion of persecution invariably depends on homosexuality.” (p. 266) In short, Freud explicitly allowed that if the young woman is paranoid, then her case is a refuting instance of the aetiology he had postulated for that disorder. Alternatively, he reckoned with the possibility that she was not paranoid.

As it turned out, during a second session the patient’s report on episodes at her place of employment not only greatly enhanced the likelihood of her being afflicted by delusions, but also accorded with the postulated aetiology by revealing a conflict-ridden homosexual attachment to an elderly woman there. But the point is that the psychoanalytic aetiology of paranoia is empirically falsifiable (disconfirmable) and that Freud explicitly recognized it. For, as we saw, this hypothesis states that a homosexual psychic conflict is causally necessary for the affliction. And empirical indicators can bespeak the absence of homosexual conflict as well as the presence of paranoid delusions.
Hence this example has an important general moral: Whenever empirical indicators can warrant the absence of a certain theoretical pathogen P as well as a differential diagnosis of the presence of a certain theoretical neurosis N, then an aetiologic hypothesis of the strong form "P is causally necessary for N" is clearly empirically falsifiable. It will be falsified by any victim of N who had not been subjected to P. For the hypothesis predicts that anyone not so subjected will be spared the miseries of N, a prediction having significant prophylactic import. Equivalently, the hypothesis retrodicts that any instance of N was also a case of P. Hence, if there are empirical indicators as well for the presence of P, then this retrodiction can be empirically instantiated by a person who instantiates both N and P.

Being a strict determinist, Freud’s aetiological quest was for universal hypotheses. (S.E. 1915, 14:265) But he believed he had empirical grounds for holding that the development of a disorder N after an individual I suffers a pathogenic experience P depended on I’s hereditary vulnerability. Hence his universal aetiologic hypotheses typically asserted that exposure to P is causally necessary for the development of N, not that it is causally sufficient.

Indeed, by claiming that P is the “specific” pathogen of N, he was asserting not only that P is causally necessary for N, but also that P is never, or hardly ever, an aetiologic factor in the pathogenesis of any other nosologically distinct syndrome. (S.E. 1895, 3:135-139) Robert Koch’s specific aetiology of tuberculosis, i.e., the pathogenic tubercle bacillus, served as a model. (S.E. 1895, 3:137) By the same token, Freud pointed to the tubercle bacillus to illustrate that a pathogen can be quite explanatory, although its mere presence does not guarantee the occurrence of the illness. (S.E. 1896, 3:209) And Freud was wont to conjecture specific aetiologies for the various psychoneuroses until late in his career. (S.E. 1925, 20:55) Hence, as illustrated by the above example of paranoia, these aetiologies evidently have a high degree of empirical falsifiability, whenever empirical indicators can attest a differential diagnosis of N, as well as the absence of P. For the hypothesis that P is the specific pathogen of N entails the universal prediction that every case of non-P will remain a non-N, and equivalently, the universal retrodiction that any N suffered P, although it does not predict whether a given exposure to P will issue in N. Thus Glymour’s account (1974) of Freud’s case history of the Rat Man makes clear how Freud’s specific aetiology of the Rat Man’s obsession was falsified by means of disconfirming the retrodiction that Freud had based on it.
Let us return to our paranoia example. As I pointed out in an earlier paper (Grüna 1979a, pp. 138-139), the aetiology of paranoia postulated by psychoanalysis likewise makes an important "statistical" prediction that qualifies as "risky" with respect to any rival "background" theory that denies the aetiologic relevance of repressed homosexuality for paranoia. And, by Popper's standards, the failure of this prediction would count against Freud's aetiology, whereas its success would corroborate it.

To be specific, Freud originally hypothesized the aetiology of male paranoia (Schreber case) along the following lines. (SE. 1911, 12:63) Given the social taboo on male homosexuality, the failure to repress homosexual impulses may well issue in feelings of severe anxiety and guilt. And the latter anxiety could then be eliminated by converting the love emotion "I love him" into its opposite "I hate him," a type of transformation that Freud labeled "reaction formation." Thus the pattern of reaction formation is that once a dangerous impulse has been largely repressed, it surfaces in the guise of a far more acceptable contrary feeling, a conversion that therefore serves as a defense against the anxiety associated with the underlying dangerous impulse. When the defense of reaction formation proves insufficient to alleviate the anxiety, however, the afflicted party may resort to the further defensive maneuver of "projection" in which "I hate him" is converted into "He hates me." This final stage of the employment of defenses is then the full-blown paranoia. Thus this rather epigrammatic formulation depicts reaction formation and projection as the repressed defense mechanisms that are actuated by the postulated specific pathogen of homosexuality. But if repressed homosexuality is indeed the specific aetiologic factor in paranoia, then the decline of the taboo on homosexuality in our society should be accompanied by a decreased incidence of (male) paranoia. And, by the same token, there ought to have been relatively less paranoia in those ancient societies in which male homosexuality was condoned or even sanctioned. For the reduction of strong anxiety and repression with respect to homosexual feelings would contribute to the removal of Freud's conditio sine qua non for this syndrome.

Incidentally, as Freud explains (S.E. 1915, 14:265), before he enunciated universally that homosexuality is the specific pathogen of paranoia, he had declared more cautiously in his earlier publication that it is "perhaps an invariable" aetiologic factor. (S.E. 1911, 12:59-60 and 62-63, esp. p. 59) When I first drew the above "statistical" prediction from Freud's aetiology (Grüna 1979a, p. 139), I allowed for Freud's more cautious early formulation. For there I predicated the forecast of decreased incidence as a
concomitant of taboo decline on the *ceteris paribus* clause that no other potential causes of paranoia become operative. But, by making repressed homosexuality the *conditio sine qua non* of the syndrome, Freud’s specific aetiology clearly enables the prediction to go through *without* any such proviso.

On the other hand, even assertions of pathogenic causal relevance that are logically *weaker* than the specific aetiologies can be empirically disconfirmable. For they can have testable (disconfirmable) predictive import, although they fall short of declaring P to be causally necessary for N. Thus, when pertinent empirical data fail to bear out the prediction that P positively affects the incidence of N, they bespeak the causal *irrelevance* of P to N. Consequently the currently hypothesized causal relevance of heavy cigarette smoking to lung cancer and cardiovascular disease is disconfirmable. And so is the alleged causal relevance of laetril to cancer remission, which was reportedly discredited by recent findings in the United States.

The aetiology that Freud conjectured for one of his female homosexual patients furnishes a useful case in point. He states its substance as follows:

> It was just when the girl was experiencing the revival of her infantile Oedipus complex at puberty that she suffered her great disappoint-ment. She became keenly conscious of the wish to have a child, and a male one; that what she desired was her father’s child and an image of him. her consciousness was not allowed to know. And what happened next? It was not she who bore the child, but her unconsciously hated rival, her mother. Furiously resentful and embittered, she turned away from her father and from men altogether. After this first great reverse she forswore her womanhood and sought another goal for her libido. (S.E. 1920, 18:157)

But later on he cautions:

> We do not, therefore, mean to maintain that every girl who experiences a disappointment such as this of the longing for love that springs from the Oedipus attitude at puberty will necessarily on that account fall a victim to homosexuality. On the contrary, other kinds of reaction to this trauma are undoubtedly commoner. (S.E. 1920, 18:168)

Thus he is disclaiming the predictability of lesbianism from the stated pubescent disappointment *in any one given case*. Yet the frustration does have disconfirmable predictive import, although its causal relevance is not claimed to be that of a specific pathogen. For by designating the stated sort
of disappointment as an aetiologic factor for lesbianism, Freud is claiming that occurrences of such disappointment positively affect the incidence of lesbianism.

This predictive consequence should be borne in mind, since Freud’s case history of his lesbian patient occasioned his general observation that the aetiologic explanation of an already existing instance of a disorder is usually not matched by the predictability of the syndrome in any one given case. (S. E. 1920, 18:167-68) And an apologist for Popper was thereby led to conclude that the limitation on predictability in psychoanalysis thus avowed by Freud is tantamount to generic nonpredictability and hence to nondisconfirmability. But, oddly enough, this apologist is not inclined to regard the causal relevance of heavy smoking to cardiovascular disease as wholly nonpredictive or nondisconfirmable, although chain smoking is not even held to be a specific pathogen for this disease, let alone a universal predictor of it.

The comments I made so far in response to Popper’s aforecited 1974 statement focused largely on Freud’s 1915 paper on paranoia, whose very title announces an instance of empirical falsifiability. Besides, Freud’s 1933 “Revision of the Theory of Dreams” presents an acknowledged falsification by the recurrent dreams of war neurotics. These are mere illustrations of falsifiability. But I trust they have prepared the ground for my claim that Popper’s refutability criterion is too insensitive to reveal the genuinely egregious epistemic defects that indeed bedevil the clinical psychoanalytic method and the aetiologies based upon it. And, as I shall argue, time-honored inductivist canons for the validation of causal claims do have the capability to exhibit these cognitive deficits.

IV. Does Neo-Baconian Inductivism Sanction the Testability of Psychoanalytic Aetiologies by Clinical Data?

I should remind the reader that “clinical data” are here construed as findings coming from within the psychoanalytic treatment sessions. When I am concerned to contrast these data from the couch with observational results secured from outside the psychoanalytic interview, I shall speak of the former as “intra-clinical” for emphasis. Freud gave a cardinal epistemological defense of the psychoanalytic method of clinical investigation that seems to have gone entirely unnoticed in the literature, as far as I know, until I recently called attention to its significance. I have dubbed this
pivotal defense “The Tally Argument” in earlier publications (Grunbaum 1979b, 1980), and I shall use this designation hereafter.

It will be important for the concerns of this paper to have a statement of some of the theses for which I have argued (Grunbaum 1983, 1983b).

1. It was Freud’s “Tally Argument”—or its bold lawlike premise—that was all at once his basis for five claims, each of which is of the first importance for the legitimation of the central parts of his theory. These five claims are the following:

   (i) The denial of an irremediable epistemic contamination of clinical data by suggestion.

   (ii) The affirmation of a crucial difference, in regard to the dynamics of therapy, between psychoanalytic treatment and all rival therapies that actually operate entirely by suggestion.

   (iii) The assertion that the psychoanalytic method is able to validate its major causal claims—such as its specific sexual aetiologies of the various psychoneuroses—by essentially retrospective methods without vitiation by post hoc ergo propter hoc, and without the burdens of prospective studies employing the controls of experimental inquiries.

   (iv) The contention that favorable therapeutic outcome can be warrantedly attributed to psychoanalytic intervention without statistical comparisons pertaining to the results from untreated control groups.

   (v) The avowal that, once the patient’s motivations are no longer distorted or hidden by repressed conflicts, credence can rightly be given to his or her introspective self-observations, because these data then do supply probatively significant information. (Cf. Waelder 1962, pp. 628-29)

2. The epistemological considerations that prompted Freud to enunciate his Tally Argument make him a sophisticated scientific methodologist, quite better than is allowed by the appraisals of relatively friendly critics like Fisher and Greenberg (1977) or Glymour (1980), let alone by very severe critics like Eysenck.

Yet evidence that has been accumulating in the most recent decades makes the principal premise of the Tally Argument well-nigh empirically untenable, and thus devastatingly undermines the conclusions that Freud drew from it. Indeed, no empirically plausible alternative to that crucial discredited premise capable of yielding Freud’s desired conclusions seems to be in sight.
3. Without a viable replacement for Freud’s Tally Argument, however, there is woefully insufficient warrant to vindicate the intraclinical testability of the cardinal tenets of psychoanalysis—especially its ubiquitous causal claims—a testability espoused traditionally by analysts, and more recently by Glymour on the strength of the pincer-and-bootstrap strategy. This unfavorable conclusion is reached by the application of neo-Baconian inductivist standards whose demands for the validation of causal claims can generally not be met intraclinically, unless the psychoanalytic method is buttressed by a powerful substitute for the defunct Tally Argument. And in the absence of such a substitute, the epistemic decontamination of the bulk of the patient’s productions on the couch from the suggestive effects of the analyst’s communications appears to be quite utopian.

4. Insofar as the credentials of psychoanalytic theory are currently held to rest on clinical findings, as most of its official spokesmen would have us believe, the dearth of acceptable and probatively cogent clinical data renders these credentials quite weak. And without a viable alternative to the aborted Tally Argument having comparable scope and ambition, the future validation of Freudian theory, if any, will have to come very largely from extra-clinical findings.

5. Two years before his death, Freud invoked the consilience or convergence of clinical inductions (in the sense of William Whewell) to determine the probative cogency of the patient’s assent or dissent in response to the interpretations presented by the analyst. (S.E. 1937, 23:257-269) But such a reliance on consilience is unavailing unless and until there emerges an as yet unimagined trustworthy method for epistemically decontaminating each of the seemingly independent consilient pieces of clinical evidence. For the methodological defects of Freud’s “fundamental rule” of free association (S.E. 1923, 18:238; 1925, 20:41; 1940, 23:174) ingress alike into the interpretation of several of these prima facie independent pieces of evidence (e.g., manifest dream content, parapraxes, waking fantasies). And this multiple ingress renders the seeming consilience probatively spurious.

6. Given the aforestated dismal inductivist verdict on clinical testability, that traditional inductivist methodology of theory appraisal no more countenances the clinical validation of psychoanalysis than Popper does. (1962, p. 38, fn. 3) Hence the specifically clinical confirmations claimed by many Freudians but abjured as spurious by inductivist canons are unavail-
able as a basis for Popper’s charge of undue permissiveness against an inductivist criterion of demarcation between science and nonscience. And I have already argued in section III above that the actual falsifiability of psychoanalysis undercuts his reliance on Freud’s theory as a basis for claiming greater stringency for his criterion of demarcation.

Here I shall focus solely on the issue posed by the third of these six theses. In Grünbaum 1983a I argued in detail that there are fundamentally damaging flaws in the actual clinical arguments given by Freud and his disciples for the very foundation of his entire edifice: the theory of repression. Thus I claim to have shown there that the actual clinical evidence adduced by Freudians provides no cogent basis for the repression-aetiology of neuroses, for the compromise-model of manifest dream content, or for the causal attribution of parapraxes (“slips”) to repressed ideation. And in Grünbaum 1983, sec. III, 3 I canvassed solid evidence for the considerable epistemic contamination and hence lack of probative value on the part of three major kinds of clinical findings that Freud deemed either initially exempt from such adulteration or certifiably not marred by it because of due precautions: the products of “free” association, the patient’s assent to analytic interpretations that he (she) had initially resisted, and memories recovered from early life.

Indeed, the epistemic adulteration I documented there seems to be ineradicable in just those patient responses that are supposed to lay bare his repressions and disguised defenses after his resistances have been overcome. And yet Freud attributed pride of place to these very data in the validation of his theory of repression, a doctrine that is avowedly “the cornerstone on which the whole structure of psychoanalysis rests . . . the most essential part of it.” (S.E. 1914, 14:16) Thus, generally speaking, clinical findings—in and of themselves— forfeit the probative value that Freud had claimed for them, although their potential heuristic merits may be quite substantial. To assert that the contamination of intraclinical data is ineradicable without extensive and essential recourse to extra-clinical findings is not, of course, to declare the automatic falsity of any and every analytic interpretation which gained the patient’s assent by means of prodding from the analyst. But it is to maintain—to the great detriment of intraclinical testability!—that, in general, the epistemic devices confined to the analytic setting cannot reliably sift or decontaminate the clinical data so as to identify those that qualify as probative. And clearly all of these liabilities apply fully as much to Heinz Kohut’s “self-psychology” version of
psychoanalysis, which places the major determinants of adult personality structure into an even earlier phase of childhood than Freud did. (Cf. Meyers 1981; Basch 1980, chapter XI)

One must admire the strenuous and ingenious efforts made by Freud to legitimate his psychoanalytic method by arguing that it had precisely such an identifying capability. As I pointed out in Grünbaum 1983, these efforts included the attempt to vouchsafe the probity of free associations by excluding their contents in the bastion of internal causal relatedness. And Freud’s dialectical exertions culminated in the generic underwriting of clinical investigations by the Tally Argument. But the empirical untenability of the cardinal premise of the Tally Argument that I have documented elsewhere (Grünbaum 1980, sec. II) has issued in the latter’s collapse, leaving intraclinical validation defenseless against all of the sceptical inroads from the substantial evidence for the distortion and tailoring of its data by such mechanisms as suggestion.

Hence it is unavailing to take contaminated findings from the psychoanalytic interview more or less at face value, and then to try to employ them probatively in some testing strategy whose formal structure is rational enough as such. All the more so, since no viable substitute for the Tally Argument appears to be in sight. Indeed, the seeming ineradicability of epistemic contamination in the clinical data adduced as support for the cornerstones of the psychoanalytic edifice may reasonably be presumed to doom any prospects for the cogent intraclinical testing of the major tenets espoused by Freud.

These considerations can now be brought to bear on the scrutiny of Glymour’s defense of clinical testability, which was outlined in our Section I above.

Glymour gives an illuminating reconstruction of Freud’s account of the Rat Man case by means of the logical pincer-and-bootstrap strategy that Glymour had teased out of that account. I have no reason to gainsay this strategy in general, as far as it goes. But I shall now argue that—with or without it—strong reasons militate against the intraclinical testability of the specific aetiologic hypothesis that was at issue in the case of the Rat Man, Paul Lorenz, who suffered from an obsessional fear of rats.

At the time of the Rat Man case, Freud had postulated that premature sexual activity such as excessive masturbation, subjected to severe repression, is the specific cause of obsessional neurosis. As will be recalled from Section III, in his carefully defined usage of “specific cause,” the claim that
X is the specific cause of Y entails unilaterally that X is causally necessary for Y. And the latter, in turn, unilaterally entails that all cases of Y were X's. Thus if this particular consequence of the conjectured sexual aetiology is to get confirmation from Lorenz's psychoanalysis, the intraclinical data yielded by it need to be able to certify the following: Lorenz, who was an adult victim of obsessional neurosis, engaged in precocious sexual activity, which was then repressed. Hence let us inquire first whether intraclinical data produced by the adult patient can reliably attest the actual occurrence of a childhood event of the stated sort. But, as I shall argue, even if the answer to this question were positive, this much would be quite insufficient to support Freud's aetiologic hypothesis that repressed precocious sexual activity is causally relevant to adult obsessional neurosis.

As Glymour (1980) notes, "Freud had . . . arrived at a retrodicted state of affairs, namely, the patient's having been punished by his father for masturbation." (p. 272) And indeed, "the crucial question is whether or not Lorenz was in fact punished by his father for masturbation." (p. 273) But Freud's specific aetiology of adult obsessional neurosis as such calls only for an early childhood event in which precocious sexual activity was repressed. Why then should it be probatively "crucial" whether it was the patient's father who was involved in the sexual event required by the hypothesized aetiology?

As is clear from Freud's account, the elder Lorenz's involvement became probatively weighty, because of the unconscious significance that psychoanalytic theory assigns to the patient's recollection of recurring fears of his father's death, at least after the age of six. While having these fears, the child Paul bore his father deep conscious affection. Freud derived the presumed unconscious origin of the fears from a theoretical postulate of so-called precise contrariety, which he took pains to communicate to the patient, who then became "much agitated at this and very incredulous." (S.E 1909, 10:180) Freud explains both his reasoning and revealingly relates his indoctrination of the patient:

. . . he was quite certain that his father's death could never have been an object of his desire but only of his fear.—After his forcible enunciation of these words I thought it advisable to bring a fresh piece of theory to his notice. According to psycho-analytic theory, I told him, every fear corresponded to a former wish which was now repressed; we were therefore obliged to believe the exact contrary of what he had asserted. This would also fit in with another theoretical requirement, namely, that the unconscious must be the precise
Having thus theoretically inferred the patient's deep childhood grudge against his father from the recurring fears of losing the father, Freud also conjectured that the grudge remained so durably unconscious only because it was a response to the father's interference with the patient's sensual gratification.

And this conclusion was, then, serendipitous in suggesting that there had been an early event satisfying the specific aetiology that Freud had hypothesized for Lorenz's obsessional neurosis. Since this aetiology required precocious masturbation events, Freud retrodicted that the patient had been punished by his father for masturbation "in his very early childhood . . . before he had reached the age of six." (S.E. 1909, 10:183)

Clearly the actual occurrence of an event having these attributes would simultaneously satisfy the initial condition of the postulated aetiology and explain Lorenz's early dread of his father's death via Freud's principle of precise contrariety.

Let us now suppose, just for argument's sake, that Freud's avowedly well-coached adult patient had actually reported having a memory of the very early childhood event that Freud had retrodicted. Then I ask: Could such a clinical event have reliably attested the actual occurrence of the distant event? I have framed this question hypothetically, because it so happened that Lorenz actually had no direct memory of any physical punishment by his father, let alone of a punishment for a sexual offense. He did remember having been told repeatedly by his mother that there had been one incident of angry conflict with his father at age three or four, when he was beaten by him. And when the mother was consulted whether this beating had been provoked by a misdeed of a sexual nature, her answer was negative. Furthermore, this was apparently the only beating the child had ever received from the father.

But for the purpose of our inquiry, we are positing that, at some point in his analysis, the patient had claimed to remember just the kind of early childhood event that Freud had retrodicted via his specific aetiology of obsessional neurosis. Then I am concerned to show that, taken by itself,
such a finding would be quite insufficient to lend any significant support to the hypothesized aetiology of obsessional neurosis. And my reasons for this claim will then enable me to argue quite generally for the following conclusion: Given the demise of the Tally Argument, the intraclinical testing of the causal assertions made by Freud's specific aetiologies of the psychoneuroses, and by his ontogenetic developmental hypotheses, is epistemically quite hopeless.

Let "N" (neurosis) denote a psychoneurosis such as the syndrome of obsessional neurosis, while "P" (pathogen) denotes the kind of sex-related antecedent event that Freud postulated to be the specific cause of N. Thus I shall say that a person who had a sexual experience of the sort P "is a P." And if that person was then afflicted by N, I shall say that he was both a P and an N, or just "a PN." It is taken for granted, of course, that there are both N's and non-N's, as well as P's and non-P's. To support Freud's aetiologic hypothesis that P is causally necessary for N, evidence must be produced that being a P makes a difference to being an N. But such causal relevance is not attested by mere instances of N that were P's, i.e., by patients who are both P's and N's. For even a large number of such cases does not preclude that just as many non-P's would also become N's, if followed in a horizontal study from childhood onward! Thus instances of N that were P's may just happen to have been P's, whereas being a P has no aetiologic role at all in becoming an N.

A telling, sobering illustration of this moral is given by the following conclusion from a review of forty years of research (Frank, 1965):

No factors were found in the parent-child interaction of schizophrenics, neurotics, or those with behavior disorders which could be identified as unique to them or which could distinguish one group from the other, or any of the groups from the families of the [normal] controls (p. 191).

Hence it is insufficient evidence for causal relevance that any N who turns out to have been a P does instantiate the retrodiction "All N's were P's," which is entailed by Freud's specific aetiology. Thus, to provide evidence for the causal relevance claimed by Freud, we need to combine instances of N that were P's with instances of non-P who are non-N's. And indeed, since he deemed P to be causally necessary for N—rather than just causally relevant—his aetiology requires that the class of non-P's should not contain any N's whatever, whereas the class of P's is to have a positive (though numerically unspecified) incidence of N's.
One can grant that since “All N’s are (were) P’s” is logically equivalent to “All non-P’s are (will be) non-N’s,” any case of an N who was a P will support the latter to whatever extent it supports the former. But this fact is unavailing to the support of Freud’s aetiology. For the issue is not merely to provide evidential support for “All non-P’s are (will be) non-N’s,” or for its logical equivalent, by some instances or other. Instead the issue is to furnish evidential support for the (strong kind of) causal relevance claimed by Freud. And, for the reasons I have given, the fulfillment of that requirement demands that there be cases of non-P’s that are non-N’s no less than instances of N’s that were P’s. Yet at best, the Rat Man could furnish only the latter kind of instance. In other words, if we are to avoid committing the fallacy of post hoc ergo propter hoc, we cannot be content with instances of N’s that were P’s, no matter how numerous. Analogously, suppose it were hypothesized that drinking coffee is causally relevant to overcoming the common cold. And consider the case of a recovered cold-sufferer who retrospectively turns out to have been drinking coffee while still afflicted with the cold. Then such an instance, taken by itself, would hardly qualify as supportive of the hypothesized causal relevance.

Psychoanalytic theory and therapy have encouraged the disregard and even flouting of the elementary safeguards against the pitfalls of causal inference familiar since the days of Francis Bacon, not to speak of J. S. Mill. Yet even informed laymen in our culture are aware that such safeguards are indeed heeded in medicine before there is public assent to the validity of such aetiologic claims as “heavy tobacco smoking causes cardiovascular disease.” This double standard of evidential rigor in the validation of aetiologic hypotheses even makes itself felt in current criminal law. For legal prohibitions—and so-called expert psychiatric testimony in courts of law—are sometimes predicated on such hypotheses even when their credentials are no better than that blithe repetition has turned them into articles of faith. The recently publicized reiteration of the purported pathogenicity of child molestation in opposition to its decriminalization is a case in point.

In our society, such sexual molestation is often alleged to be pathogenic, even when it is affectionate and tender rather than violent. And this allegation has been invoked to justify making it illegal and fraught with substantial penalties. Yet recently a number of sexologists have maintained that very young children should be allowed, and perhaps even encouraged, to have sex with adults, unencumbered by interference from the law. In
their view, such activity itself is harmless to the child and becomes harmful only when parents raise a fuss about it. Indeed, some of these advocates have made the daring and quite unfashionable aetiological claim that unless children do have early sex, their psychological development will go awry. And even the less daring champions of harmlessness are opposed to jailing affectionate pedophiles.

Reasons of elementary prudence and also of humaneness make it a good policy, in my view, to put the burden of proof on those who maintain that affectionate and tender child molestation is not even distressing to the child, let alone pathogenic. But a cautionary basis for a legal prohibition is a far cry from the confident assertion of demonstrated pathogenicity. And the difference between mere caution and authenticated causation of neurosis may, of course, be relevant to the severity of the punishment appropriate for violations of the interdiction.

In a recent issue of *Time* magazine, John Leo inveighs aetologically against the demand to legalize tender pedophilia, which he sees as a thinly disguised manifesto for child molesters liberation. And the justification offered by him for his indictment is as follows:

Unfortunately, few responsible child experts have reacted . . . so far to the radical writing on child sex. One who has is Child Psychiatrist Leon Eisenberg of Children’s Hospital Medical Center, Boston: “Premature sexual behavior among children in this society almost always leads to psychological difficulties because you have a child acting out behavior for which he is not cognitively or emotionally ready. . . .” Psychotherapist Sam Janus, author of a new book, *The Death of Innocence,* says that people who were seduced early in life “go through the motions of living and may seem all right, but they are damaged. I see these people year after year in therapy.” U.C.L.A. Psychiatrist Edward Ritvo also says that much of his work is with children who have been involved in catastrophic sexual situations. His conclusion: “Childhood sexuality is like playing with a loaded gun.” (September 7, 1981, p. 69)

But the aetiological reasoning of those whom *Time* cites to document the pernicious effects of child molestation is just as shoddy as the causal inferences of those advocates of pedophilia who claim dire psychological consequences from the failure of infant boys to act on their erections, and of infant girls to utilize their vaginal lubrications. For the findings adduced by *Time* do not answer either of the following two questions:
(1) Is the occurrence of childhood seduction not equally frequent among those who are well enough never to see a psychotherapist? In the parlance of John Stuart Mill, this question calls for the use of the joint method of agreement and difference, rather than just the heuristic method of agreement.

(2) Would a representative sample of those who were not seduced in childhood have a significantly lower incidence of adult neurosis than those who were seduced? By the same token, we must ask those who claim seduction to be beneficial psychologically to show that those who were indeed seduced fared better psychologically than those who were not sexually active in this way. Without the appropriate answers to these questions, the respective assertions of causal relevance remain gratuitous. Thus we must ask those who condemn childhood seduction the foregoing questions, because it may be that childhood seduction just happens to be quite common among neurotics, even though it has no aetiologic role in the production of neurosis. In that case, the same people would have become neurotics anyway, in the absence of early seduction. Without answers to these questions, the evidence given by those whom Time invokes as authorities merely suggests the bare possibility that childhood seduction is pathogenic. By the same token, psychoanalysts have overlooked the fact that repressed homosexual feelings cannot be shown to be the pathogen of adult paranoia, by merely pointing to the frequency of homosexually-tinged themes in the associative output of paranoiacs during analysis. This finding does not tell us whether homosexual themes would not likewise turn up to the same extent in the so-called free associations of nonparanoiacs who lead well-adjusted lives and never see a therapist. Here no less than in the case of the Rat Man, the invocation of J. S. Mill’s heuristic method of agreement is not enough to lend support to the hypothesis of aetiologic relevance.

Hence even if the Rat Man did in fact have the sexually repressive experience P retrodicted via Freud’s aetiology of obsessional neurosis, this alone would hardly qualify as evidential support for that aetiology. And there is a further reason for concluding that even if the child Paul Lorenz had actually been punished by his father for masturbating, as retrodicted via Freud’s aetiology, this putative occurrence would confer little, if any, support on this aetiology. For, as Ronald Giere has remarked (in a private communication), the occurrence of this sort of event is to be routinely
expected in the Victorian child-rearing culture of the time on grounds other than psychoanalytic theory.

Moreover, Freud had made the adult Rat Man patient well aware, as we saw, of the inferences that Freud had drawn about his childhood via psychoanalytic theory. Given the substantial evidence I adduced for the notorious docility of patients in analysis (Grünbaum 1983), I submit that one ought to discount Lorenz's putative early childhood memory as too contaminated to attest reliably the actual occurrence of the retrodicted early event.

It can be granted, of course, that requirements of consistency or at least overall coherence do afford the analyst some check on what the patient alleges to be bona fide memories. But Freud's own writings attest to the untrustworthiness of purported adult memories of early childhood episodes that had presumably been repressed in the interim and then retrieved by the analysis. (Cf. the documentation in Grünbaum 1980, p. 353) Indeed, the malleability of adult memories from childhood is epitomized by a report from Jean Piaget (Loftus 1980, pp. 119-121), who thought he remembered vividly an attempt to kidnap him from his baby carriage along the Champs Elysées. He recalled the gathered crowd, the scratches on the face of the heroic nurse who saved him, the policeman's white baton, the assailant running away. However vivid, Piaget's recollections were false. Years later the nurse confessed that she had made up the entire story, which he then internalized as a presumed experience under the influence of an authority-figure.

The discounting of the Rat Man's putative early childhood memory is hardly a general derogation of the reliability of adult memories in ordinary life. But in the clinical context, the posited memory is simply not sufficiently dependable to qualify as evidence for the retrodicted event. Thus the retrospective intra-clinical ascertainment of the actual occurrence of the retrodicted distant event is just too unreliable. And, in general, the patient's memory may simply fail to recall whether the pertinent event did occur, as Freud himself stressed. (S.E. 1920, 18:18; 1937, 23:265-66) Indeed, even in survey studies of lung cancer patients who are asked about their prior smoking habits, and of heroin addicts who are questioned about previous use of marijuana, the retrospective ascertainment of the actual occurrence of the suspected causal condition is epistemically flawed. (Giere 1979, pp. 216, 265)

Have I provided adequate grounds for maintaining that long-term
prospective studies, which employ control groups and spring the clinical confines of the usual psychoanalytic setting, must supplant the retrospective clinical testing of aetiology defended by Glymour? Not just yet. For suppose that analysts could secure reasonable numbers of patients who, though presumed to need analysis for some neurotic affliction or other, are certifiably free of the particular neurosis N (say, obsessional neurosis) whose aetiology is currently at issue. Since neuroses usually occur in mixed rather than pure form, this is a generous assumption. All the same, let us operate with it. Then if we are given such patients who, though neurotic, are non-N, Freud’s pertinent specific aetiology does not retrodict whether patients of this sort were P’s or non-P’s. For his hypothesized pathogenesis allows given non-N’s to have been P’s no less than to have been non-P’s, although it does require any non-P to become a non-N. Now postulate, for argument’s sake that, though retrospective, psychoanalytic inquiry were typically able to ascertain reliably whether a given case of non-N was indeed a non-P or a P. If so, then non-N’s who putatively turn out to have been P’s would merely be compatible with Freud’s aetiologic hypothesis instead of supporting it, since this hypothesis allows these instances without requiring them.

But what of patients who are neither N’s nor P’s? Would such people, together with other persons who are both N’s and P’s, jointly bespeak that P is pathogenic for N (obsessional neurosis) within the class of all persons?

Note that the clinical testing design I have envisaged for scrutiny is confined to the class of neurotics. For even the non-N’s of this design are presumed to be afflicted by some neurosis other than N. The reason is that persons who have practically no neuroses of any sort are hardly available to analysts in sufficient numbers to carry out the putative retrospective determination of whether they were non-P’s or P’s. But, as Mr. Blake Barley has noticed, the confinement of this retrospective clinical determination to the class of neurotics has the following consequence: Even if every observed non-N (non-obsessive neurotic) is a non-P while every observed N is a P, these combined instances lend credence only to the hypothesis that, within the class of neurotics, P is aetiologically relevant to N. But these putative combined instances do not support the Freudian claim of such aetiologic relevance within the wider class of persons.

In short, the Freudian clinical setting does not have the epistemic resources to warrant that P is neurotogenic. And this unfavorable conclu-
sion emerges even though it was granted, for argument's sake, that the retrospective methods of psychoanalytic inquiry can determine reliably whether adult neurotics who are non-obsessives were non-P's in early life.

But is it reasonable to posit such reliability? It would seem not. For clearly, even if the patient believes to have the required memories, the retrospective clinical ascertainability of whether a given non-N was actually a non-P is epistemically on a par with the psychoanalytic determination of whether a given N was a P. And, as we saw, the latter is unreliable. Moreover, as Freud himself acknowledged, “the patient cannot remember the whole of what is repressed in him, and what he cannot remember may be precisely the essential part of it.” (S.E. 1920, 18:18)

Now contrast the stated epistemic liabilities of the retrospective psychoanalytic inference that a given adult patient was or was not a P during his early childhood with the assets of prospective controlled inquiry: A present determination would be made, under suitably supervised conditions, that children in the experimental and control groups are P's and non-P's respectively; again, during long-term follow-ups, later findings as to N or non-N would also be made in what is then the present.

Recently experimental validations of therapeutic efficacy have been carried out by using the response history of single individuals without control groups drawn from other individuals. (Hersen and Barlow 1976; Kazdin 1981; Kazdin, in press) Thus in these validations, the causal claims inherent in the pertinent assertions of therapeutic efficacy have been validated by single-case experimental designs. Hence it behooves us to ask whether these “intrasubject” validations could become prototypic for using a given analysand to test intraclinically the causal assertions made by the long-term aetiologic hypotheses of psychoanalytic theory and by such claims of efficacy as are made for its avowedly slow therapy. To answer this question, let us first look at situations in physics in which the probative equivalent of controlled experiments is furnished by other means.

When a billiard ball at rest on a billiard table suddenly acquires momentum upon being hit by another billiard ball, we are confident that the acceleration of the first ball is due to the impact of the second. And even more importantly, astronomers made sound causal claims about the motions of planets, binary stars, etc. before they were able to manipulate artificial earth satellites, moon probes, or interplanetary rockets. What took the probative place of control groups in these cases? In the case of the
billiard ball, Newton’s otherwise well-supported first law of motion gives us background knowledge as to the “natural history” of an object initially at rest that is not exposed to external forces: Such an object will remain at rest. And this information, or the law of conservation of linear momentum, enables us to point the finger at the moving second billiard ball to furnish the cause of the change in the momentum of the first. A similar reliance on otherwise attested background knowledge supplies the probative equivalent of experimental controls in the astronomical cases.

Turning to the single-case validations of therapeutic efficacy, they pertain to the following sort of instance:

. . . a seven-year-old boy would beat his head when not restrained. His head was covered with scar tissue and his ears were swollen and bleeding. An extinction procedure was tried: the boy was allowed to sit in bed with no restraints and with no attention given to his self-destructive behavior. After seven days, the rate of injurious behavior decreased markedly, but in the interim the boy had engaged in over ten-thousand such acts, thus making the therapists fearful for his safety. A punishment procedure was subsequently introduced in the form of one-second electric shocks. In a brief time, the shock treatment dramatically decreased the unwanted behavior. (Erwin 1978, pp. 11-12)

Here the dismal prospects of an untreated autistic child are presumably known from the natural history of other such children. And, in the light of this presumed background knowledge, the dramatic and substantial behavior change ensuing shortly after electric shock allowed the attribution of the change to the shock without control groups. For, under the circumstances, the operation of other causal agencies seems very unlikely. More generally, the paradigmatic example of an intrasubject clinical validation of the causal efficacy of a given intervention is furnished by the following variant of using the single patient as his own “historical” control: (i) The natural history of the disorder is presumably otherwise known, or (ii) the therapist intervenes only in on-off fashion, and this intermittent intervention is found to yield alternating remissions and relapses with dramatic rapidity.

Can the causal validation designs employed in these intrasubject clinical tests of therapeutic efficacy become prototypic for using an individual analysand to validate Freud’s long-term aetiologic hypotheses, or to furnish evidence that an analysis whose typical duration extends over several years deserves credit for any therapeutic gain registered by the
patient after, say, four years? To ask the question is to answer it negatively. The natural history of a person not subjected to the experiences deemed pathogenic by Freudian theory is notoriously unknown! And as for crediting therapeutic gain to analytic intervention on the basis of an intrasubject case history, how could such an attribution possibly be made in the face of Freud’s own acknowledgment of the occurrence of spontaneous remissions (S.E. 1926, 20:154)? At best, Freudians can hope to show that the incidence of therapeutic gain in groups of patients who undergo analysis exceeds the spontaneous remission rate in untreated control groups belonging to the same nosologic category (Rachman and Wilson 1980). In short, the stated intrasubject validation by means of dramatic therapeutic gains can hardly be extrapolated to underwrite the prospective single-case evaluation of slow analytic therapy, let alone to vindicate the retrospective testing of a Freudian aetiology in the course of an individual analysis.

Although Freud’s specific aetiologies did not specify numerically the percentage of P’s who become N’s, it is noteworthy that only prospective investigation can yield the information needed for such a statistical refinement. For let us suppose that retrospective data confirm the retrodiction of Freud’s specific aetiology that the incidence of P’s within the sample group of N’s is 100 percent. Then this percentage incidence clearly does not permit an inference as to the percentage incidence of N’s within the class of P’s. Yet such information is clearly desirable, if only in order to estimate the probability that a child who was subjected to P will become an N. More generally, when P is not deemed causally necessary for N, but merely causally relevant, retrospective data simply do not yield any estimates of P’s degree of causal effectiveness. (Giere 1979, pp. 274, 277)

Our inquiry into the Rat Man case so far has operated with a counterfac-tual posit in order to discuss the reliability of clinical data by reference to this case. The hypothetical clinical datum we used was that the patient had reported having a memory of the early childhood event retrodicted by Freud. As against Glymour’s generic thesis that the specific psychoanalytic aetiologies can be cogently tested “on the couch,” I have argued that, at least typically, such testing is epistemically quite hopeless. And hence it would seem that Paul Lorenz’s psychoanalysis would have completely failed to furnish evidential support for the aetiologic relevance of childhood sexual repression to obsessional neurosis, even if Paul’s father had reliably reported having repeatedly punished his young son for masturbation.
Incidentally, when Waelder defended the clinical confirmation of the psychoanalytic aetiologies, he overlooked precisely that their substantiation requires evidential support for the causal relevance of the purportedly pathogenic experience, and not merely the historical authentication of the bare occurrence of that experience. (1962, pp. 625-626)

Let us return to Glymour’s account of the testing strategy in Paul Lorenz’s analysis, which was predicated on Lorenz’s failure, in actual fact, to have any direct recall of receiving a punishment from his father, let alone a castigation for a sexual offense. Therefore let us now see how Glymour evaluated the probative significance of this finding. I shall be concerned to stress the scope that Glymour does give to essential reliance on extraclinical data for probative purposes. Indeed, it will turn out that the entire testing procedure in the Rat Man case comes out to be probatively parasitic on an extraclinical finding. And hence I wonder how Glymour can see himself as rebutting Eysenck’s denial of intraclinical testability, although he does succeed in impugning the demand that all extraclinical disconfirmation be experimental.

By Glymour’s own account of the Rat Man case, the probatively “crucial” data came from the extraclinical testimony of the patient’s mother. On Glymour’s reading of Freud, at the time of Lorenz’s analysis, Freud still postulated actual rather than fancied early sexual experiences to be the pathogens of obsessional neurosis. (Glymour 1980, pp. 274-275) And as Glymour explains lucidly, what made Lorenz’s case a counterexample to this aetiology was not the mere failure of the patient to recall the event retrodicted by Freud. Instead, it was the extraclinical testimony from the mother that had this negative probative import. (p. 273) For it was her testimony that supplied the probatively crucial datum by contravening Freud’s retrodiction, when she answered the question that Glymour characterized as “the crucial question.” And he himself characterizes “the memory of an adult observer”—in this case that of the mother—as “the most reliable available means” for answering this decisive question as to the character of the offense for which the child Paul had been punished. (p. 273) How, then, in the face of the extraclinical status of the decisive datum, can Glymour justify his description of the testing rationale used in the Rat Man case as “a strategy that relies almost exclusively on clinical evidence”? (Glymour 1974, p. 287)

It is true enough that, as we know from the case history of the Wolf Man, Freud regarded stories told by other members of the family to the patient
about his childhood to be generally “absolutely authentic” and hence as admissible data. (S.E. 1918, 17:14 and fn. 2) But Freud completes this assertion by cautioning that responses by relatives to pointed inquiries from the analyst—or from the patient while in analysis—may well be quite contaminated by misgivings on their part:

So it may seem tempting to take the easy course of filling up the gaps in a patient’s memory by making enquiries from the older members of his family; but I cannot advise too strongly against such a technique. Any stories that may be told by relatives in reply to enquiries and requests are at the mercy of every critical misgiving that can come into play. One invariably regrets having made oneself dependent upon such information; at the same time confidence in the analysis is shaken and a court of appeal is set up over it. Whatever can be remembered at all will anyhow come to light in the further course of analysis. (S.E. 1918, 17:14, fn. 2)

Even if one were to discount Freud’s caveat, several facts remain: (i) it is misleading to claim intraclinical testability if, as in the Rat Man case, the avowedly crucial datum does not come from “the couch.” (ii) What makes the reliance on extra-clinical devices important is that, far from being marginal epistemically, its imperativeness derives from the typically present probative defects of the analytic setting, defects that are quite insufficiently acknowledged by Glymour. And in my view, it does not lessen the liabilities of intraclinical testing that the compensations for its deficits from outside the clinical setting may occasionally be available in situ (e.g. from family records) and thus do not necessarily have to require the experimental laboratory. For even when supplemented by such nonexperimental, extra-clinical devices, the thus enlarged “clinical” testing procedure is not adequate or epistemically autonomous. For example, when it becomes necessary to resort to extraclinical information for the sort of reason that was operative in the Rat Man case, it will be a matter of mere happenstance whether suitable relatives are even available, let alone whether they can reliably supply the missing essential information. Why then dignify as a “clinical testing strategy” a procedure of inquiry that depends on such contingent good fortunes and hence, when luck runs out, cannot dispense with experimental information? (iii) The real issue is whether the clinical setting typically—rather than under contingently favorable circumstances—does have the epistemic resources for the cogent validation of the aetiology at issue in the Rat Man case, and of other analytic aetiologies. In dealing with that issue, Glymour’s otherwise illuminating
account has not demonstrated the existence of a cogent intraclinical testing strategy, even if he succeeded in showing that extraclinical compensations for its lacunae need not be wholly experimental.

Indeed, the extent of his essential epistemic reliance on extraclinical findings can now be gauged from his view of the effect that Freud's modifications of the specific sexual aetiology of obsessional neurosis (and of other neuroses) had on the testability of these evolving aetiological hypotheses. Glymour (1980) recounts this evolution:

After the turn of the century and before 1909, . . . there is no statement of the view that sexual phantasies formed in childhood or subsequently, having no real basis in fact, may themselves serve in place of sexual experiences as etiological factors. . . . Yet after the Rat Man case the view that either infantile sexual experiences or phantasies of them may equally serve as etiological factors became a standard part of Freud's theory. In Totem and Taboo, four years after the Rat Man case appeared, Freud emphasized that the guilt that obsessional neurotics feel is guilt over a happening that is physically real but need not actually have occurred [fn. omitted]. By 1917 Freud not only listed phantasies themselves as etiological factors alternative to real childhood sexual experiences, but omitted even the claim that the former are usually or probably based on the latter [fn. omitted]. The effect of these changes is to remove counterexamples like that posed by the Rat Man case, but at the cost of making the theory less easily testable. For whereas Freud's theories, until about 1909, required quite definite events to take place in the childhood of a neurotic, events that could be witnessed and later recounted by adults, Freud's later theory required no more than psychological events in childhood, events that might well remain utterly private. (pp. 276-27)

Thus Glymour attributes the diminishing testability of Freud's modified aetiologies quite rightly to the lessening extraclinical ascertainability of the sorts of events that Freud successively postulated as aetiologic. But if the testability of the psychoanalytic aetiologies is in fact "almost exclusively" intraclinical, as Glymour told us, why should it be vital for their testability that the aetiologic events required by Freud's later theory are just mental states of the patient to which only the patient himself and his analyst become privy?

Incidentally, the problem of testing Freud's sexual aetiology of the neuroses—either clinically or extraclinically—became less well defined after he gave up the quest for qualitatively specific pathogens of nosologic-
cally distinct psychoneuroses in favor of a generic Oedipal aetiology for all of them. In fact, he used the analogy of explaining the great qualitative differences among the chemical substances by means of quantitative variations in the proportions in which the same elements were combined. But having thus dissolved his prior long-standing concern with the problem of "the choice of neurosis," he was content to leave it at vague metapsychological remarks about the constant intertransformation of "narcissistic libido" and "object libido." (S.E. 1925, 20:55-56)

What of Glymour's reliance on *intraclinical* data? In that context, he seems to have taken much too little cognizance of even the evidence furnished by analysts that intraclinically the suggestibility problem is radically unsolved, if not altogether insoluble because there is no viable substitute for the defunct Tally Argument. Can we place any stock in Glymour's aforecited aspiration that clinicans can be "trained so as not to elicit by suggestion the expected responses from their patients"? In view of the evidence for the ineradicability of suggestive contamination, it would now seem that this hope is sanguine to the point of being quite utopian. In an Afterword to his (1974) for a second edition of the Wollheim volume in which it first appeared, Glymour (1982) has reacted to some of these particular doubts as follows:

I do not see . . . that the experimental knowledge we now have about suggestibility requires us to renounce clinical evidence altogether. Indeed, I can imagine circumstances in which clinical evidence might have considerable force: when, for example, the clinical proceedings show no evident sign of indoctrination, leading the patient, and the like; when the results obtained fall into a regular and apparently law-like pattern obtained independently by many clinicians; and when those results are contrary to the expectation and belief of the clinician. I do not intend these as *criteria* for using clinical evidence, but only as indications of features which, in combination, give weight to such evidence.

To this I say the following:

1) I do *not* maintain that any and all clinical data are altogether irrelevant probatively. Instead, I hold that such findings cannot possibly bear the probative burden placed upon them by those who claim, as Glymour did, that psychoanalysis can TYPICALLY be validated or invalidated "on the couch," using a clinical testing strategy that is mainly confined to the analytic setting.
(2) The existence of some circumstances under which we would be warranted in not renouncing clinical evidence "altogether" is surely not enough to sustain clinical testing as a largely cogent and essentially autonomous scientific enterprise. And as for Glymour's illustrations of such circumstances, I cannot see that absence of evident indoctrination, or regular concordance among the results obtained independently by many clinicians exemplify circumstances under which "clinical evidence might have considerable force." For—apart from the arguments I gave against these illustrations, if only à propos of "free" association (Grübaum 1983)—it seems to me that their utopian character as a step toward solving the compliance problem is epitomized by the following sobering results, which are reported by the analyst Marmor (1962):

... depending upon the point of view of the analyst, the patients of each [rival psychoanalytic] school seem to bring up precisely the kind of phenomenological data which confirm the theories and interpretations of their analysts! Thus each theory tends to be self-validating. Freudians elicit material about the Oedipus Complex and castration anxiety, Jungians about archetypes, Rankians about separation anxiety, Adlerians about masculine strivings and feelings of inferiority, Horneyites about idealized images, Sullivanians about disturbed interpersonal relationships, etc." (p. 289)

(3) I do not deny at all that now and then clinical results "are contrary to the expectations and belief of the clinician." But as a step toward vindicating clinical inquiry qua epistemically autonomous testing strategy, I can only say "One swallow does not a summer make."

What seems to me to emerge from Glymour's interesting reconstruction is that, on the whole, data from the couch acquire probative significance when they are independently corroborated by extraclinical findings, or are inductively consilient with such findings in Whewell's sense. Thus I do not maintain that any and all clinical data are altogether irrelevant probatively. But this much only conditionally confers potential relevance on intraclinical results beyond their heuristic value. And surely this is not enough to vindicate testability on the couch in the sense claimed by its Freudian exponents, and countenanced by Glymour.

Note

Adolf Grünbaum

Press, 1955), 10:155-318. This paper first appeared in 1909. Hereafter any references given to Freud's writings in English will be to this Standard Edition under its acronym "S. E." followed by the year of first appearance, the volume number, and the page(s). Thus the 1909 paper just cited in full would be cited within the text in abbreviated fashion as follows: S. E. 1909, 10:155-318.

References


