The Journal of Mind and Behavior
Summer 1999, Volume 20, Number 3
Pages 311–350
ISSN 0271–0137

A Close and Critical Examination of How Psychopharmacotherapy Research is Conducted

David H. Jacobs

California Institute for Human Science

This paper conducts a critical examination of the "usual psychopharmacology standards" for clinical research. Four main areas are inspected: (a) What is "it" that is being treated in a clinical drug study? (b) How much is really known concerning the psychological alterations brought about by psychiatric drug treatment? (c) To what extent are sources of bias actually controlled for in "controlled" drug treatment studies? (d) How does the usual "dropout pattern" influence the alleged "clinical findings" of controlled drug treatment studies? The overall conclusion reached is that the "usual standards" cannot produce a realistic picture of either safety or efficacy. Conceptual and methodological reforms are suggested.

The present paper is devoted to a critical examination of the "usual psychopharmacology standards" for clinical research. I have borrowed this term from the 1989 NIMH investigation of short-term treatments for depression (Elkin et al., p. 978), but I intend to go well beyond what this NIMH research team identified as problematic. Specifically, I will examine the following "usual" features of psychopharmacotherapy research:

1. The background claim that the clinical problem at hand belongs to the realm of biomedicine, since it is this claim which seems to so "naturally" justify the basic reality of psychopharmacotherapy research, that is, physicians administering drugs to patients in a medical setting. A related feature of psychopharmacotherapy research, perhaps not so readily recognized as part of a conventional biomedical framework of thought, will likewise be examined: the claim that discrete (separate and independent) psychiatric disorders really exist and can be clinically treated *sui generis*.

Requests for reprints should be sent to David H. Jacobs, Ph.D., Psychotherapy Associates, 528 Fourth Street, Encinitas, California 92024.

2. Controlled clinical trials are primarily designed for the purpose of establishing the *efficacy* of the investigational drug compared to placebo (and perhaps another drug) in the short-term treatment of a specific psychiatric condition (diagnosis). Much less thought and effort is directed at the problem of establishing the full range of psychological alterations brought about by what is obviously a psychoactive substance. In fact, virtually all of the researchers' interest in "side effects" in formal, hypothesis-testing clinical studies is directed at either somatic distress or what could be called the lowest level of drug-induced psychological disturbances (restlessness, agitation, nervousness, etc.). Drug effects in the realm of psychosocial functioning are ruled out by default in controlled studies in virtue of non-investigation. The problem of relying (almost) exclusively on the drugged person's own ability to detect drug-induced psychological decrements or distortions is unrecognized and unaddressed methodologically. The upshot is that (a) the full range and severity of drug treatment "costs" cannot be brought into decision-making about drug treatment, and (b) the fundamental question of whether drug-induced symptom reduction is actually salutary/therapeutic or part of a detrimental drug-induced altered state is decided in favor of the first possibility by default.

3. "Controlled" drug treatment studies fail to control for multiple sources of bias which may substantially influence findings. In addition to the obvious fact that the usual "blinding" arrangement fails to accomplish its purpose, a range of additional sources of bias are identified, most of which stem from the brevity of controlled studies.

4. The solicitation of research subjects on the basis of (free) drug treatment inevitably results in a larger number of placebo-treated subjects terminating treatment prematurely than subjects who are treated with the active medication. The inability of *brief* drug treatment trials to keep what may be a majority of placebo-treated subjects in treatment for the full planned course of the trial results in an incomplete comparison between drug efficacy and placebo efficacy. This is unacknowledged by research reports which advance what is called "end-point analysis" of the clinical data as the "real" findings of the investigation. The problem of differential dropout rate between drug treatment and placebo treatment is set up by the usual reluctance of researchers to incorporate the comparison which should be made to pharmacological treatment, namely various forms of psychotherapy.

Mystification and Misdiagnosis

On what could be characterized as the eve of psychiatry's official abandonment of the psychosocial framework of thought — I am referring of course to the 1980 publication of the DSM-III, with all its attendant fanfare — Gerald

Klerman (1978) published an explicit list of "tenets" upon which neoKraepelinism (remedicalization) rested.[1] Of primary importance were the principles that psychiatric conditions were to be regarded as bona fide medical illnesses in the conventional medical sense, and that psychiatric illnesses were to be regarded as discrete, mutually independent clinical entities (diseases, disorders, syndromes, what have you), again in the conventional medical sense. To make what Klerman meant in the latter case clear, if it is not, I can draw upon an earlier paper by Robins and Guze (1970) which has often been cited as a landmark in psychiatry's eventual recommitment to diagnosis within a medical framework. Under the heading "delimitation from other disorders," Robins and Guze observe that entirely different diseases may share certain symptoms, for example (their example) cough and blood in the sputum in lobar pneumonia, bronchiectasis, and bronchogenic carcinoma. Nevertheless, these are separate, distinct, and independent diseases, and this can be established on the basis of entirely different causal pathogenic agents, pathophysiology, prognosis, course, and so on. In point of fact Klerman was (in the 1978 paper) unable to bring forward evidence, as he admitted, that any one of the usually designated *functional* psychiatric illnesses actually was a medical condition in the conventional sense, or even a discrete and autonomous clinical entity, but that was not the point. The point was that it was necessary to accept the neoKraepelinian tenets in order for psychiatry to move ahead, especially with regard to biological research. The tenets, in other words, served the purpose of redefining psychiatry as a field of biomedical practice and research. As such, they were not themselves in need of proof.

What I wish to bring out by the foregoing observations is that psychopharmacotherapy research since the neoKraepelinian revolution rests upon two paradigm-defining claims which are taken for granted, but which cannot bear the test of evidential challenge. These two claims constitute the *background understanding* out of which drug treatment studies are conducted. Were I to simply pass over these background *tenets* without challenge, then my more focused criticisms concerning methodology and data analysis would lose much of their force. The background tenets, again, are these: first, what is being treated in a psychopharmacotherapy investigation is a bona fide medical disease, more specifically a form of neuropathology. The experimental subject's suffering is *not* akin to, say, the despair a person may feel when his/her business is failing and going into bankruptcy, or the demoralization a person may feel under the cumulative weight of many adverse life events experienced over time. In a 1980 publication, Klerman made it clear that a

[1]From the 1970s until his death in 1992, Gerald Klerman was one of the leading figures in the remedicalization of psychiatry. He appears to have had a rather substantial change of heart shortly before his death. I provide more discussion on this point later in this section.

psychiatric diagnosis was distinct from being unhappy and so on as a consequence of "social deprivation or the frustration of personal wishes" (Weissman and Klerman, 1980). This is, in other words, only a different semantic version of the tenet that psychiatric conditions are genuine *medical* illnesses, not merely psychological reactions (even if severe) to the vicissitudes of life. In 1996, the new Director of NIMH, Steven E. Hyman, M.D., in his first address as the new Director to readers of *Psychotherapy and Rehabilitation Research Bulletin* (an NIH–NIMH publication), reiterated the position that "mental disorders are medically valid illnesses that affect the brain. There can be no doubt that mental illnesses reflect disorder within the brain . . ." (p. 1). There certainly is no doubt that this *position* has been consistently presented to the public as a non-disputable scientific fact. The truth-value of this position is another matter, as I will discuss. There also can be little doubt that public opinion has been markedly influenced in the direction of accepting the neoKraepelinian tenet that psychiatric conditions are actual brain disorders. For example, David Karp's (1993, 1996) ethnographic research with people who had a substantial history of treatment with so-called antidepressant drugs found that they all sooner or later adopted the view that they suffered from a "biochemical disorder."

The second *tenet* upon which psychopharmacotherapy studies rest is that what the experimental subjects are being treated for in the investigation is a genuine clinical entity. That is, a legitimate and valid diagnosis has been made and is the basis for the treatment the subjects are receiving. So, for example, the 1989 NIMH study concerning short-term treatments for depression (Elkin et al.) purports to treat a specific clinical entity/diagnostic category, namely "major depressive disorder." The authors list a number of additional psychiatric diagnoses which, if detected in potential research subjects, would exclude them from the study. No explanation is provided as to why these diagnoses and not others should exclude subjects. Presumably some of the diagnoses (schizophrenia, panic disorder) were selected as exclusion criteria because it could be argued that persons who are simultaneously ill in these other respects should be treated first for their even more urgent problems. Nevertheless, only a few diagnostic categories are named as exclusion criteria (only one from the entire class of personality disorders, namely antisocial personality disorder), so it is clear that the vast majority of DSM-III-R disorders may be copresent without excluding the candidate and without being thought by researchers to demand immediate treatment (since concurrent treatment for any additional condition is another exclusion criterion). The tenet here is that the accepted research subjects have a specific disorder which can and should be treated *sui generis*. I will show that this tenet is a fiction, a fiction which is *not* in the best interests of the subjects.

The more basic issue is whether people who complain of depression, anxiety, panic, etc. are medically ill. Contemporary biopsychiatry in effect asserts the existence of neurological illnesses which even over the course of decades manifest themselves *only* by limited forms of disturbance — that is, over time *no* somatic indications of neuropathology appear and *no* indications of mental deterioration (*not* just peculiarities) appear: no tremor, chorea, seizures, paralysis, dyskinesia, deterioration of psychomotor skills, balance, coordination, manifest losses in sensation, perception, deterioration in memory, speech, reasoning, abstraction, learning, comprehension, etc. (M. Bleuler, 1978, 1991; Carson and Sanislaw, 1993; Dumont, 1984; Heinrichs, 1993; Kendell, 1991; McHugh, 1991; Sarbin, 1990). The situation indeed appears fundamentally paradoxical: researchers believe that what they are treating is a form of neuropathology, but if signs of neuropathology are actually observed in a potential research subject, the person is eliminated from experimental study on the grounds of presenting signs of neuropathology. For example, the 1989 NIMH study of short term treatments for depression (Elkin et al.) excluded potential subjects who exhibited signs of organic brain syndrome (a broad, non-specific reference to neuropathology). As M. Bleuler (1978) pointed out in his longitudinal study of schizophrenia, the fundamental issue in medicine is to determine whether a real biomedical condition is present or not. This does not in any way suggest that the person's complaints of depression, fear, fatigue, pain, etc. should be dismissed or denigrated if no biological disorder can be detected. But in the absence of objective signs of biological disorder, the physician should refrain from declaring that a genuine medical illness is definitely present. With regard to traditional medical thinking concerning schizophrenia, Bleuler laments that this fundamental principle of medical reasoning has been ignored, and physicians who point out that objective signs of physical illness are lacking in schizophrenia are subject to ridicule. Nevertheless, it should be observed that nothing is more ambiguous in medicine than psychological disturbance (emotional distress, etc.) that is unaccompanied by objective signs of biological disorder in the present, and which over time does not progress to include objective signs of biological disorder (see also Ananth, 1984; Caine and Shoulson, 1983; Hoffman, 1982; Rodin and Voshart, 1986; Strub and Black, 1985; Taylor, 1982; Warnes, 1982).

Since people with a psychiatric diagnosis do not routinely progress to manifest signs of physical illness, on what grounds could it nevertheless be claimed that a psychiatric condition is brought about by "medically valid illness," specifically "disorder within the brain" (Hyman, 1996)? It would be necessary to show that the right kind of neurological lesion exists in people who qualify for psychiatric diagnosis.[2] The right kind of lesion would have

[2]By "lesion" I mean a detectable injury or dysfunction in the brain.

the following form: (i) it would distinguish between psychiatric patients and other neurological patients. This is necessary because the (putative) psychiatric neurological lesion must be compatible with the fact that psychiatric patients typically do not progress to visible somatic illness and mental deterioration over time; (ii) it would show the correct temporal relationship to symptoms, that is, a detectable lesion [which fulfills requirement number (i)] would be shown to *cause* the relevant behavioral symptoms, not merely *co-exist with* the relevant behavioral symptoms (for example, heightened white blood cell count is the result of infection, not its cause); (iii) it would distinguish between people who are clinically ill and people who are not clinically ill. This is necessary because a "lesion" derives its meaning from manifest illness, not merely from the fact that it is unusual. Without this requirement, it would be possible to "show" that a person is psychiatrically ill in the absence of behavioral symptoms — a paradoxical outcome (Cochrane, 1972; Duster, 1984; Eisenberg, 1980; Engel, 1977; Gevirtz et al., 1996; Ross, 1986, 1995). Does evidence exist to support the kind of neurological illness that is advanced by contemporary psychiatry[3] — demonstrated neuropathology that produces *only* limited psychological symptomatology (and not deterioration)? If we look for physical evidence of any kind which fulfills the three criteria listed above (or any of the criteria considered separately), and which in addition has been shown over time to be replicable across samples of research subjects and across research teams and sites — in short, evidence recognized by the scientific community as real evidence — there is simply nothing at which to point (Andreason and Carpenter, 1993; Bogerts, 1993; Carson and Sanislaw, 1993; Heinrichs, 1993; Janicak et al., 1993; Kendell, 1991; Kirch, 1993; Lieberman and Koreen, 1993; Mesulam, 1990; Pam, 1990; Ross, 1995; Shear et al., 1993). To say that there is biological research which looks "promising" is simply to repeat what has been said in psychiatry throughout

[3]A question might be raised concerning just how broad and inclusive contemporary medically-oriented psychiatrists view the scope of neuropathological explanation in psychiatry. This is a difficult question to answer with certainty. Note that NIMH Director Steve Hyman (1996) equates mental disorder with disorder within the brain without qualification in his first official communique in *Psychotherapy and Rehabilitation Bulletin* (an NIH–HIMH publication). The DSM-IV only identifies two disorders as definitely brought about (that is, caused) by virtue of noxious external circumstances, namely PTSD and Adjustment Disorder. But it is revealing that the DSM-IV specifically *prohibits* a diagnosis of Adjustment Disorder, *regardless* of external circumstances, if the individual qualifies symptomatically for an Axis I diagnosis (p. 623). In a 1993 paper concerning contemporary progress in diagnosing schizophrenia, Andreason and Carpenter point out that the ultimate goal of medical diagnosis is definition of a disease in terms of etiologic mechanisms and pathophysiology. This would seem to make it clear that they have categorically rejected psychogenesis as an appropriate *form* and *level* of explanation for schizophrenia. Since their comments concerning the ultimate goal of medical diagnosis appear under a general discussion of diagnosis in psychiatry, it seems legitimate to conclude that they reject psychogenesis as explanatory far beyond schizophrenia itself.

the entire course of this century — and always in the same spirit, which is to dismiss as irrelevant the patient's history of mistreatment, trauma, abuse, neglect, victimization, and so on, no matter how severe and chronic.

What of the argument that the *organic* nature of psychiatric illness is revealed by the fact that people who are depressed and so on can be helped with medication? The most obvious objection to this argument is the well-established finding that entirely non-somatic forms of treatment are at least as efficacious as drug treatment (Antonuccio et al., 1995; M. Bleuler, 1978, 1991; Carpenter et al., 1977; Ciompi, 1997; Easton and Link, 1986; Greenberg and Fisher, 1989; Karon, 1989; Marks et al., 1993; Mavissakalian and Michelson, 1986; Menn and Mosher, 1982; Milrod, 1993; Milrod et al., 1996; Mosher and Burti, 1989; Shea et al., 1992; Shear et al., 1993; Vaughn and Leff, 1976; Vega and Murphy, 1990; Wing, 1987). In addition, the Elkin et al. (1989) "NIMH Treatment of Depression Collaborative Research Program" study found that imipramine (the "standard" for antidepression drug treatment efficacy) did no better than placebo in relieving depression over a 16 week period. This finding is particularly significant because the Elkin et al. study stands practically alone in taking measures to create at least one set of clinical observations which were well protected from researcher bias. The implication is that a large portion of the "benefit" usually found in psychopharmacotherapy research derives from the ability of researchers to detect which subjects are taking the active medication despite the "blinding" procedure (see ahead).

If the claim can be dismissed that psychiatric illness is really a form of neuropathology, what of the discrete syndromes (diagnostic categories like major depressive disorder) for which people are treated in psychopharmacotherapy studies? In general medicine a syndrome denotes a distinctive "clinical entity," that is, a distinctive set of signs and symptoms (synchronically) and a distinctive course and prognosis (diachronically). Medical knowledge concerning etiology and pathophysiology are not necessary for syndromal identification. What is at issue is whether a distinct clinical entity can be identified on the basis of symptomatology, course, and prognosis. It could hardly be more obvious that in psychiatry the signs and symptoms of disturbance or distress that real patients "present" to clinicians show no respect at all for discrete diagnostic categories (syndromes). The categories, in short, are *definitional* rather than *empirical*, they come into existence by committee either in virtue of the discrete illness tenet itself, or (like "major depressive disorder") in virtue of the intention to coordinate certain signs and/or symptoms with drugs which are widely thought or hoped to be useful in ameliorating just these signs and/or symptoms. It is in fact rarely the case that the person who receives a DSM diagnosis suffers only from the defining features of the diagnosis. The "diagnosis" then, and in contrast to what a "diagnosis"

means in general medicine, does *not* answer the question "What is the matter with this patient? — What should be treated?" In short, a psychiatric diagnosis is not at all what is understood as a diagnosis in general medicine. Published discussions on this matter reveal that the non-category and non-diagnostic status of DSM-III type diagnosis is entirely appreciated by the psychiatric elite (Andreason and Carpenter, 1993; Frances et al., 1991; Shear et al., 1993; Strauss, 1986, 1992; Strauss et al., 1979, 1985; Terr, 1991; Vaillant, 1984, 1988; Widiger and Shea, 1991).[4]

What is the point of defining syndromes which do not empirically exist? I believe the most fundamental answer is this: biopsychiatry collapses without discrete syndromes. Emotional distress, cognitive peculiarities, and behavioral disturbances *must* fall naturally into discrete syndromes with predictable courses or there is *no* possibility of ever showing that "psychiatric illness" can be explained in terms of pathophysiology. There must *be* a genuine syndrome in order for its etiopathogenesis to be discovered — no amount of biological research can discover the cause of a non-existent syndrome (this point of logic comes up repeatedly in both medical research and psychiatric research literature, e.g., Austin, Stolley, and Laskey, 1992; Bolos et al., 1990; Dumont, 1984; Duster, 1990; Edwards, 1991; Kendell, 1991; Lewontin, Rose, and Kamin, 1984; Lidz, Blatt, and Cook, 1981; Marshall, 1994; Pam, 1990; Polymeropoulos et al., 1996; Ross, 1995). For example, "the cause" of schizophrenia cannot yield to biological research if in the main the actual symptoms displayed by "schizophrenics" overlap/crisscross with the actual symptoms displayed by patients given other diagnoses (e.g., "schizophrenics," patients with affective psychosis, patients with dissociative disorders, patients with borderline personality disorder, and patients with various other diagnoses can suffer from hallucinations and/or delusions). Categorical diagnosis is arbitrary when the relevant empirical data in fact indicate continuity rather than discontinuity, and it is then logically fallacious to suppose that an underlying biological order can be discovered to justify arbitrary assignments to separate (diagnostic) categories.[5] The same point was made in a 1983 publication concerning the (then) new DSM-III hyperkinetic/attention deficit syndrome by the British psychiatrist, Michael Rutter. As he put the matter, a diagnostic category must be distinctive in some way other than simply elaborating upon the symptoms that supposedly define it. Thus it is no

[4]In a revealing article concerning the development of the DSM-III based upon unpublished American Psychiatric Association archival material, Wilson (1993) cites a memo from Robert Spitzer to the DSM-III Task Force which explains that symptom clusters must be called syndromes or disorders for the purpose of insurance reimbursement.

[5]This point is made with admirable clarity by Kendell and Brockington, 1980, in the context of reporting their own failure to empirically distinguish between patients given the diagnosis of schizophrenia and patients given some other psychotic diagnosis.

use asserting that what makes (for example) attention deficit syndrome a distinct clinical entity is the presence of attention difficulties when this is in fact a common feature of numerous other diagnoses. Rutter (like Kendell and Brockington, 1980) makes it clear that the creation of a non-valid syndrome cannot sustain hope that "it" represents a coherent underlying biological reality. By contrast, Huntington's disease was identified as a syndrome long before medical technology was at a level of sophistication which made causal discoveries or practical treatment possible (Caine and Shoulson, 1983). The fact that the psychiatric symptomatology associated with Huntington's disease is variable and unpredictable did not prevent its syndromal identification precisely because order and predictability exist at the level of somatic symptoms (chorea, etc.) and course. But in psychiatry there is no other orderly level — there is only subjective distress and cognitive and behavioral peculiarities which do not as a matter of fact arrange themselves naturally (i.e., empirically) into syndromes, considering both symptomatology at time of initial clinical contact and course over time.

My previous remark to the effect that a DSM "syndrome" is actually a statement of intent to treat some signs of distress/disturbance rather than others is primarily meant to bear on the relationship between Axis I disorders and pharmacotherapy. In a 1991 publication on the topic of conceptual problems facing the DSM-IV, Task Force members Widiger and Shea grope to formulate a reason for the original distinction made in DSM-III between Axis I disorders and Axis II disorders. It is clear that they are unable to find a theoretically persuasive reason for the original creation of the Axis I–Axis II distinction (or for retaining it in DSM-IV), but they do suggest in passing that "responsivity to pharmacologic treatment" may have been one of the guiding principles. They further admit that splitting-off various islands of blatant symptomatology (e.g., depressed mood) from disorders of personality is "problematic" and even "illusory." But they do not draw-out the clinical consequences of designating islands of symptomatology as distinct and independent disorders (diagnostic categories) and acting *as if* such split-off conceptual creations can and should be treated *sui generis*.

A thoughtful discussion of this matter can be found in what would appear to be an unlikely source. I am referring to a 1993 paper on the topic of reintroducing a psychodynamic understanding of the origin and treatment of panic disorder by Shear, Cooper, Klerman, Busch, and Shapiro. The first author, Katherine Shear, served on the DSM-III-R committee on anxiety disorders. In 1988 she was still of the opinion that panic disorder was primarily a neurobiological illness (Shear and Fryer, 1988). I have already introduced the name Gerald Klerman and given some indication of his role in the remedicalization of American psychiatry during the 1970s and beyond. Klerman's contribution to the Shear et al. paper can hardly be appreciated

without knowledge that during the 1980s he co-directed the Upjohn-sponsored multinational clinical trials for FDA Approval of Xanax as a treatment for panic disorder. In a 1988 paper which offered an overview of the entire project, Klerman made it clear that panic disorder was considered a discrete diagnostic entity and a neurobiological illness. To find both Shear and Klerman on the author list of the 1993 paper must be seen as a remarkable turnabout only five years after the 1988 papers.

The 1993 paper has two major components. One component consists of a review of clinical research on the origin and development of panic episodes, treatment possibilities, and follow-up studies. The second component consists of a report of findings of a one hour psychodynamic interview with nine newly diagnosed panic disorder patients at the Payne Whitney Anxiety Disorders Clinic.

In the component devoted to review of clinical research literature, the authors make the following points: the postulated neurobiological disturbance in panic disorder has failed to materialize despite continued research; people diagnosed with panic disorder in fact present a history of broad spectrum psychological difficulties as well as a broad spectrum of contemporaneous psychological difficulties; and *successfully* treated panic disorder patients nevertheless continue to suffer from numerous psychological difficulties and impairments. In other words, a contemporary diagnosis of panic disorder is indicative of a more broadly troubled past and present, and the "successfully" treated person with panic disorder (whether treated pharmacologically or with a form of focused, brief psychotherapy) remains ill, disabled, impaired, etc. Since it appears to be the case that a diagnosis of panic disorder does not adequately capture "what is the matter" with the person who is given the diagnosis, and since the *successfully* treated panic disorder patient is still quite disturbed and remains so over time, the authors cite with approval the conclusion of a group of Australian clinical researchers (Andrews et al., 1990) that realistic treatment must go beyond immediate symptomatology to address "vulnerability factors," for example, immature defenses and methods of coping. Having endorsed such a conclusion, it is not surprising that Shear et al. state that the overarching point of their work is to reintroduce a psychodynamic perspective for understanding the development of panic attacks and for treating the vulnerabilities which predispose to attacks.

The research component of the paper compares clinical information concerning nine consecutive "panic disorder" patients based upon a standard structured interview and an initial psychodynamic interview. In each case, the panic disorder diagnosis was arrived at by employing one of two widely accepted research instruments: the Structured Clinical Interview for DSM-III-R (SCID) or the Anxiety Disorders Interview Schedule — Revised. Additional diagnoses arrived at in the foregoing manner for each patient, if

any, were also provided in a table (Table 1, p. 860) by Shear et al. for the reader's information. Additional diagnoses (four of the nine cases) were of course limited to other Axis I disorders, since the structured diagnostic interviews which were used for assessment purposes do not cover Axis II disorders.

Shear et al. do not describe the one hour psychodynamic interview beyond the comment that it "was similar to the usual first interview in a psychoanalytically oriented office practice" (p. 860). The interviews were videotaped and detailed written transcripts were made. Then all the transcripts were discussed by the group of five researchers for the purpose of identifying patient problems, conflicts, ways of coping, defenses, and so on. It is apparent that Shear et al. do not expect the psychodynamic interviews to simply reproduce the diagnoses arrived at on the basis of the structured research interviews (the SCID or the Anxiety Disorders Interview Schedule — Revised), nor did this happen. What may not be as apparent is that the authors are de facto making a definite statement concerning the clinical utility and validity of the structured interview method for arriving at a diagnosis, or even for generating useful clinical information at all. So, for example, whereas the SCID essentially places the patient in the position of answering a pre-set series of questions, preferably by indicating yes or no, a "usual" psychodynamic interview is tantamount to the opposite of this, that is, the patient is permitted and encouraged to bring up and embellish upon his/her own concerns in his/her own way. What emerges from these two methods is largely incommensurable, a fact that both sides realize very well.[6] Indeed, one way of explaining why formal, categorical diagnosis was regarded with such indifference under the old psychosocial paradigm is to point out that when the patient is permitted to express him/herself freely, the feasibility of a pre-established category system into which the patient can readily be sorted disappears (this point is covered well in Wilson, 1993).

The contrast between the diagnostic picture of the nine subjects based upon the structured interview compared to the picture based upon — I emphasize this point — the single psychodynamic interview, is quite dramatic. I emphasize the fact that only a single, first contact psychodynamic interview was conducted to remind the reader that first contact diagnosis is required institutional policy, reimbursement policy, and/or research requirement — not because it can be argued that first contact information is adequate for realistic diagnosis. Indeed, clinical studies which follow patients over substantial periods reveal how inadequate initial diagnostic formulations generally tend to be (Kendell, 1988; Wallerstein, 1986). This is why

[6]What is incommensurable is this: structured interviews are designed to identify the presence of one or more simply defined discrete disorders ("caseness"), while the psychodynamic interview is designed to encourage/permit the emergence of highly complex and individualized psychological difficulties.

Strauss (1992) has referred to DSM-III-type diagnosis, based upon current obvious symptomatology, as amounting to little more than "cross sectional hairsplitting" devoid of prognostic implications. The Shear et al. study is an indication of how unrealistic it is to attempt to circumvent the time and effort required to formulate a reasonably valid answer to the diagnostic question, namely "What is the matter? What should be treated?" The point I emphasize here is that psychopharmacotherapy research depends upon a picture of the subject/patient that *goes along with* what is expected of the investigational drug, but this picture falls far short of realistic diagnosis.

What is Really Known About the Psychological Alterations Brought on by Psychiatric Drugs?

It is taken for granted in psychopharmacotherapy discourse that clinically effective dosages and durations of all drugs produce at least some unwanted consequences. Obviously a realistic overview of the desirability of drug treatment (benefits vs. costs) cannot be formulated if the drawbacks side of the picture is not adequately filled in. In this section I present two lines of dissent against the conventional view that the costs of psychiatric drug treatment are in the main known and understood, and in the main, acceptable.

The first line of dissent concerns the blatant disparity which exists between side effects established in randomized, placebo-controlled clinical trials (RCTs) versus the much broader range and severity of adverse drug reaction reports which emanate from non-RCT formats. It is not clear what accounts for this disparity, but I will suggest that researcher bias and discretion make important contributions.

The second line of dissent emphasizes that RCTs are designed primarily for the purpose of establishing the clinical *efficacy* of a drug versus placebo (and perhaps another drug) in the short-term treatment of a specific condition (Axis I diagnosis). Much less attention, effort, and methodological ingenuity is directed at the problem of how to discern the presence of subtle drug-induced psychological alterations. Indeed, it could be observed that this is *de facto* regarded as a non-problem in conventionally conducted psychopharmacotherapy RCTs. Nothing in the design of such studies could be said to address the inherent complexities involved in obtaining a realistic view of the psychological alterations which may be brought about by a psychoactive drug. Not only is the subject not asked about drug effects in the realm of psychosocial functioning, but in addition the entire problem of drug-induced *anosognosia* (impairment in self-awareness and self-monitoring) — which should explicitly be recognized as *the* fundamental source of ambiguity in conducting clinical research with psychoactive drugs — is rarely if ever men-

tioned in the method and discussion sections of psychopharmacotherapy research reports. Yet from the perspective of clinical neurology (e.g., Fisher, 1989) anosognosia has been characterized as "one of the general rules of cerebral dysfunction." If psychoactive drugs accomplish nothing else (clinically speaking), they can be counted on to disturb normal CNS physiology (Hyman and Nestler, 1996). The second line of dissent, then, focuses on the ambiguity of symptom reduction/alleviation when this is brought about by psychoactive drug treatment. The ambiguity to which I refer concerns the meaning of drug-induced symptom reduction — is drug-induced symptom reduction a salutary/therapeutic effect of the drug, or is it only a sort of illusion which depends upon not recognizing that the drug has brought about an altered state which is incompatible with normal emotional life (as in frontal lobe syndrome, for example)? This sort of ambiguity can only be addressed by investigations which recognize the complexity of discerning the full complement of psychological alterations which may be produced by prolonged exposure to a psychoactive drug, and which incorporate relevant design features into investigation. The concern that biological interventions in psychiatry may be conterminous with iatrogenic damage, impairment, dysfunction, etc. has a long history (Clark, 1956; Cole, 1960), but among contemporary North American psychiatrists only Peter Breggin has consistently advanced the proposition that the so-called therapeutic effects of psychiatric drugs are simply misunderstood features of the toxic/impairing consequences of drug treatment (e.g., 1983, 1991, 1993, 1997). The present exposition will show, at least, that conventionally conducted psychopharmacotherapy RCTs are not designed to detect subtle adverse psychological drug effects.

How Much Danger? Thoughts on the Disparity Between What is Seen in RCTs vs. Non-RCTs

Adverse drug reaction reports which derive from "open" (not an RCT format) clinical trials at university-based research and treatment centers and from "the field" (day to day clinical practice) are published in numerous psychiatry journals. Psychopharmacotherapy researchers and drug-treatment oriented clinicians are uncertain how to regard this continuous stream of clinical observation which *prima facie* indicates that the range and severity of adverse drug reactions established for psychiatric medications in RCTs is far from complete. If a medication (e.g., Prozac) has been studied in an RCT format at numerous sites and by numerous independent research teams over time, the question of why even a low frequency adverse drug reaction has escaped recognition in previous RCTs obviously comes to mind. The fact that the adverse drug reaction being reported has not already been identified in previous RCTs and does not emerge from a new RCT is, I surmise, gener-

ally seen to reflect negatively on the probative value of an adverse drug reaction report. A recent review of fluoxetine (Gram, 1994) can be taken as illustrating conventional thinking in the psychopharmacotherapy research community on this matter. In Table 3 (side effects and safety, p. 1358) and in the accompanying text, Gram distinguishes between side effects which have been established in RCTs and "suspected (rare) adverse drug reactions" which are based upon case reports. The implication is that alleged adverse drug reactions which derive from case reports are probably very uncommon *or* are not veridical drug effects at all (i.e., coincidence).

Notwithstanding the importance of the RCT format (this includes researcher and subject "blinding") for establishing statistical evidence that a drug is in and of itself a source of clinical improvement, certain forms of case reports provide compelling evidence that it is the treatment drug that is responsible for the appearance of a new feature of psychopathology advanced as an adverse drug reaction.[7] I will draw upon an adverse fluoxetine reaction report emanating from the Yale Child Study Center (King et al., 1991) in order to show that case reports may provide convincing evidence that what is being proposed as an adverse drug reaction is in all likelihood a drug effect. King et al. report that between April 1, 1988 and November 15, 1990, six children (age eight to seventeen) of 42 (14%) who had been treated with fluoxetine for various Axis I diagnoses developed *de novo* or dramatically intensified self-injurious ideation or behavior. The question is, have the authors observed a genuine drug effect or not (would the so-called adverse drug reactions have emerged in these six children without fluoxetine treatment, presumably as a consequence of how their psychopathology was developing)? For example, R., a 14-year-old female being treated with fluoxetine for obsessive–compulsive disorder (OCD) [she was also diagnosed, simultaneously, with dysthymia secondary to OCD and possible overanxious disorder], not only made a suicide attempt after five months on fluoxetine but in hospital on 40 mg/day of fluoxetine began pulling out her hair, slamming her legs and hands into objects, and burned herself with a lighter. These novel behaviors could of course be only coincidentally related to fluoxetine treatment. R.'s self-injurious behavior (and accompanying mental state, however this could or should be described) first emerged on fluoxetine, abated when fluoxetine was discontinued, emerged again following renewed fluoxetine treatment, and once again cleared when fluoxetine was discontinued. R. was one of two or perhaps three of the 6 patients described who exhibited this pattern, which can be represented schematically as in Figure 1.

[7]If no one who was diagnosed as major depressive disorder ever improved without fluoxetine, there would be no need for an RCT. For example, no one with scurvy improves without vitamin C, and continued absence of vitamin C leads to further deterioration and death. But for most medical conditions the situation is hardly so clear cut, least of all in psychiatry.

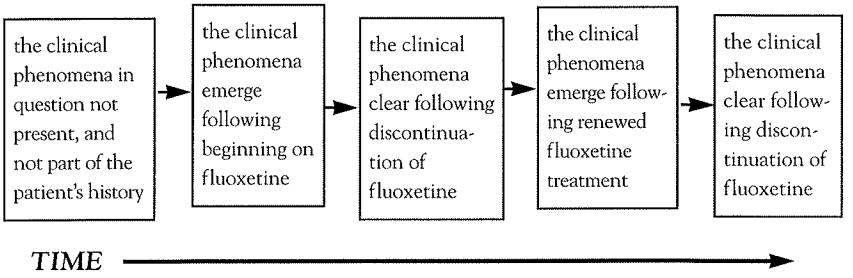

Figure 1: Relationship between fluoxetine use and discontinuance and emergence and clearing of self-injurious behavior for two or three patients reported by King et al., 1991.

In the discussion section, King et al. take note of the fact that in at least two of the 6 cases reported the proposed adverse drug reaction followed the pattern schematized above. Nevertheless, they insist that the issue (real drug effect or not?) remains in doubt without data from untreated or alternately treated contrast groups. King et al. appear to be saying that if some patients in a similarly mixed diagnosis group who do not receive fluoxetine also deteriorate and exhibit comparable symptomatology, then the question of whether the present patients really exhibited an adverse fluoxetine reaction could be answered in the negative. This line of reasoning does not seem defensible. Some patients in a similarly mixed diagnosis group may deteriorate and become self-injurious for *other* reasons. Should this occur, fluoxetine is by no means therefore exonerated or ruled-out. The question is why did *these* fluoxetine-treated patients develop the pathological behaviors observed? For the two (or perhaps three) cases in which a second trial on fluoxetine produced similar consequences it is difficult to avoid concluding that fluoxetine treatment brought about (was the proximal cause of) self-injurious behavior. King et al. speculate that fluoxetine may have brought about self-injurious behavior due to some features of its multifaceted psychoactive effect (agitation, disorganization, excitation, etc.). Note that Teicher, Glod, and Cole (1993) make a similar point in response to some of the critical comments evoked by their 1990 adverse fluoxetine reaction report. They point out that it was not their intent to suggest that fluoxetine (a chemical substance) could induce a specific thought or impulse (suicide), but rather that it *interferes with normal thought processes* so as to bring about a drug-induced obsessive–compulsive state. The comments provided by King et al. and by Teicher et al. both explicitly acknowledge that fluoxetine is a psychoactive drug which can bring about diverse and variable psychological alterations. Indeed, King et al.'s final remark should serve as one of the foundation principles in the field of psychopharmacotherapy: "Like all psy-

chotropic agents, the behavioral and neuropharmacological effects of fluoxetine are complex and variable" (p. 185).

It strikes me as peculiar that King et al. question whether fluoxetine was the proximal cause of *de novo* self-injurious behavior in at least the two cases in which the drug-contingent on–off–on–off pattern emerged because — with regard to individuals — this sort of evidence appears definitive. Of course it would be even more conclusive to show that the on–off drug-contingent sequence could be extended further, but for obvious reasons this is not feasible. The practice of "rechallenge" is fairly common in psychiatric drug treatment and research. Whatever one thinks about the ethics of rechallenge, it does appear to provide conclusive evidence that the drug is the cause of *x* clinical phenomenon.

But the typical adverse drug reaction report (Figure 1) published in psychiatry journals describes only half the rechallenge pattern, as shown in Figure 2.

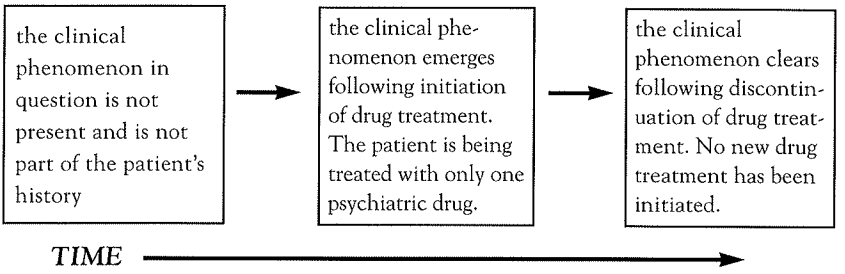

Figure 2: Typical temporal relationship between drug treatment and cessation and emergence and clearing of putative adverse drug reactions noted in adverse drug reaction reports.

In the 1990 Teicher et al. adverse fluoxetine reaction report to which I have referred, only one (case one) of the 6 cases described in the report fits the form schematized directly above. The other reported cases are complicated by multiple simultaneous drug treatment. In the 1993 paper by the same authors a new case (case seven) is introduced. This case also fits the form schematized above, but with the additional (and even more convincing) features that the person being treated was not a psychiatric patient, had no psychiatric diagnosis, had never before been treated with psychotropic drugs, and was not a "substance abuser." She was treated with fluoxetine in the hope that it would to some extent provide relief for a four-year history of chronic fatigue disorder. After 31 days of fluoxetine treatment she developed obsessive, all-consuming thoughts about killing people she loved as well as thoughts of killing herself. Although she discontinued fluoxetine (20 mg/day) on day 31, her condition as described above persisted for six months. At six months norfluoxetine (a metabolite of fluoxetine) was still detectable

in her blood. This case obviously stands apart from the *apologia* (King et al., 1991) that a psychopathological condition may take a sudden turn for the worse independently of an adverse drug reaction.

In his 1994 review paper on fluoxetine Gram notes that "rare serious drug reactions with fluoxetine is to a large extent based on case reports, and the incidence rates are unknown" (p. 1358). It is true that most adverse drug reaction reports do not provide information which would allow an estimate as to incidence, although obviously this is by no means uniformly the case. Gram's conclusion that the "suspected" adverse fluoxetine reactions he cites in Table 3 are rare is evidently based upon the fact that the reactions have not been recognized in RCTs. I have never seen an adverse drug reaction report in a psychiatry journal which questions why the adverse drug reaction being reported has not already been noticed in prior RCTs.

How can the disparity between "side effects" identified in RCTs and adverse drug reactions detected by researchers and clinicians outside the RCT format be understood? I cannot answer this question with certainty, but the following considerations are relevant. First, an RCT implicates a major financial and professional stake in the drug. Second, researchers have considerable discretion when it comes to delegating a complaint (on the part of the subject) or an observation (on the researcher's part) to the category "consistent with the subject's psychiatric condition" or to the category "drug reaction." Third, researchers have considerable leeway in naming a drug reaction. An illustration of considerations two and three can be found in a well-known 1989 adverse fluoxetine reaction report by Lipinski et al. (not cited in Gram's 1994 summary paper). The report is based upon open clinical trials with fluoxetine at the Mailman Research Center, McLean Hospital, Belmont, Massachusetts, conducted prior to FDA approval of fluoxetine. Lipinski et al. admit that the first case of fluoxetine-induced akathisia was not recognized until well into the trial by a nurse, and it was only from then onward that the researchers began to see akathisia *as such* in patients. The senior authors do not discuss how they characterized the same behaviors and complaints on the part of patients before they realized that fluoxetine could produce akathisia, but the switch from not recognizing akathisia to recognizing akathisia in their patients illustrates researcher discretion when it comes to recognizing and naming adverse drug reactions. Fourth, the natural bias which exists in favor of an experimental treatment drug is inadequately controlled in conventionally conducted RCTs. There is little reason to believe that *procedurally* blinded researchers are blind *in fact*. Indeed, it is to my mind baffling how the research community has been able to convince itself that the conventional double-blind arrangement achieves its intended aim. For example, Cohn and Wilcox (1985) describe their double-blind study of fluoxetine, imipramine, and placebo treatment for major depressive disorder

as including dosage adjustment based upon weekly interviews concerning effectiveness and side effects. This means that some fluoxetine-treated patients will soon complain about a variety of distressing states which began after fluoxetine initiation: dizziness, sedation, tremor, nausea, nervousness, insomnia, asthenia, etc. The treating psychiatrists, who it goes without saying are thoroughly familiar with the effects of fluoxetine, adjust these patients' dosages accordingly, and yet are still presumably blind as to the drug status of *all* patients.[8]

Some forms of adverse drug reaction reports (i.e., "rechallenge" reports) have more credibility than data from an RCT format concerning a specific individual's drug reaction. Naturally this conclusion, if taken seriously, would have profound effects upon how physicians think about the risks involved in treating emotional distress pharmacologically. However, adverse drug reaction reports rarely directly address Breggin's contention that a psychiatric drug's alleged therapeutic effect is actually a misinterpreted/misidentified feature of the drug's total deleterious impact on the person. The remainder of this section addresses the question of whether psychoactive medications have therapeutic properties that are separate and independent of their toxic/adverse effects.

The Ambiguity of "Doing Better" on a Psychiatric Drug: A Clinical Illustration and Its Implications for the "Ecological Validity" of RCTs

The question of what "feeling less depressed" *means* when this is brought about by a psychoactive drug is directly raised as problematic in an adverse Prozac reaction report by Hoehn–Saric, Lipsey, and McLeod (1990) of The Johns Hopkins University School of Medicine. Three patients treated for depression with Prozac are depicted as becoming euthymic (normal mood) as well as developing symptoms which clinically resemble frontal lobe lesions: apathy, flatness of affect, lack of emotional concern, loss of motivation and initiative, and difficulty foreseeing the outcome of an action.[9] As reported, the frontal lobe lesion-like symptoms were not in evidence before the initiation of fluoxetine treatment, and they remitted following discontinuation of fluoxetine.

---

[8]A much fuller discussion of the non-control of bias in psychopharmacotherapy RCTs is provided in the next section of this paper.

[9]This report emanates from clinical practice, not clinical research, which usually means — as in this case — that the clinicians providing the report do not indicate the prevalence of the adverse drug reaction in their practice. Presumably most adverse drug reaction reports which emanate from clinical practice represent the clinician's surprise at detecting a drug reaction which is not covered in the relevant medical literature. In other words, the report is a communique about something unexpected, not a summary of a planned study.

It is highly instructive to focus on just *how* Hoehn–Saric et al. eventually recognized Prozac's frontal lobe dysfunction-like effects in their patients. I am basing my remarks upon the brief clinical summaries the authors provide. Throughout the remainder of this section I will direct attention to the *de facto* assumption in psychopharmacotherapy RCTs that adverse psychological drug reactions can be detected by the researchers despite the following conditions: the researcher and subject are strangers, the "speech situation" for the subject is highly restricted, and the researchers' opportunity to observe the subject as a complex psychological being is highly restricted. Fortunately, what Hoehn–Saric et al.'s patients said in their clinical contacts with the authors was not restricted to a research protocol: this means that the patients were free to provide spontaneous commentary about how they were faring on Prozac. The picture presented by Hoehn–Saric et al. is that they slowly revised their view of Prozac's effects as their information base expanded over time. It could be said that over time Hoehn–Saric et al. began to see these patients from the perspective of clinical neurology, that is, in terms of drug-induced neuropathology, rather than in terms of medicine-induced clinical improvement. The fact that the clinical data which instigated this shift in perspective were slow to emerge and could not immediately be recognized (as signs of neuropathology) is not at all unusual in clinical neurology. Since Hoehn–Saric et al. cite Stuss and Benson's 1986 text on the frontal lobes in the context of diagnosing frontal lobe lesion-like symptoms in their patients, I will add that Stuss and Benson conclude that diagnosing frontal lobe dysfunction often requires a focused process of individualized case study. This is necessitated by the frequently subtle and distinctly non-uniform nature of frontal lobe dysfunction, which renders standardized psychometric testing more or less useless. The obvious point here is that an individualized case study approach to the question of drug-induced neuropathology is beyond the methodological scope and interest of conventional psychopharmacotherapy RCTs. For example, Hoehn–Saric et al. report that one of their patients, a fifty-year-old woman who worked as an illustrator, eventually complained that she was only able to complete projects at work while on Prozac if she was persistently reminded and cajoled to do so by others. Obviously this is not something that the authors could directly observe, and it evidently took some time before the patient brought this to their attention. They interpreted this finally as Prozac-induced apathy.

As is typical, Hoehn–Saric et al. express no curiosity as to why the Prozac effects they observed had not been identified in RCTs (they also mention in passing having observed Prozac-induced hypomania). Nevertheless, it seems evident that Hoehn–Saric et al.'s eventual identification of Prozac-induced frontal lobe dysfunction symptomatology would hardly be noticed in the drug treatment research situation. As I pointed out, the relatively unstructured

speech situation in clinical treatment, which permits a flow of narrative about everyday life, becomes a highly structured and narrowly focused question and answer session about symptoms and side effects in the *research* situation. In an RCT neither the researcher nor the subject is supposed to know if the subject has been administered the investigational drug. The study is designed to assess drug *efficacy*, not to solicit narratives about the effects of the drug *per se*. In the clinical treatment situation the patient knows that he/she is taking a specific drug, and there is nothing to restrict the patient's interest in discussing the manifold effects of the drug. It further seems apparent from Hoehn–Saric et al.'s clinical summaries that the patients only slowly become aware of what Prozac had brought about, partly on the basis of input from people in their everyday lives. It took time for the patients to realize that their drug-induced psychological condition had become problematic. The element of time stands out in each case, and this contrasts conspicuously with the six week time frame of RCTs conducted for FDA approval. There seems to be a critical missing step in psychopharmacotherapy research. The missing step is a systematic effort to establish a full picture of the psychological alterations brought about by a psychoactive drug which may be used as medicine. Reports from the field — from medical practice — are already too late, and in any event the medical treatment situation is not adequate for the task. I will return to these points.

I turn now to Hoehn–Saric et al.'s ambivalence concerning whether symptoms of frontal lobe injury which emerged on Prozac should be regarded as toxic effects unrelated to Prozac's antidepressant action (literally, side effects), or if the drug-induced apathy etc. is an important part of what is taken to be the therapeutic (antidepressant) effect. This is hardly word play, since the authors are aware that what looks clinically like frontal lobe dysfunction is definitely psychologically and socially impairing for their patients. In other words, it is not mere word play to refer to drug effects as bringing about a condition that clinically resembles frontal lobe lesions. A direct implication of referring to the effects of Prozac in this manner is that the drug has brought about a state of affairs in which the patient is reduced as a psychosocial being. This is a serious matter, obviously, one which goes well beyond the single issue of whether or not the patient feels less depressed on the drug. Presumably this is the reason — in addition to the patient's own concerns and complaints — that the authors discontinued treatment with Prozac when it became apparent that Prozac had brought about frontal lobe dysfunction-like symptoms.

On the other hand, the authors only took two of the three patients in whom they detected Prozac-induced frontal lobe injury off Prozac. The patient they kept on Prozac, like the two others, had a long and unsatisfactory history of drug treatment prior to Prozac. She felt far less depressed on Prozac, but she

had also become impaired. The authors write that "she continued on Prozac because of the substantial improvement in her mood, but we slowly tapered the dose" (p. 345). I cannot tell if this means that she voluntarily decided to stay on Prozac or if the authors made this decision. Based upon their description, it hardly seems that she was in a state to decide anything important.[10]

Her dosage was tapered down from 40 mg/day to 20 mg four days per week and 40 mg three days per week. The problematic effects of Prozac on this regimen are described as partially improved. In this decision, presented unapologetically and without further discussion, as well as in their final comments, the authors seem to abandon the position that a desirable reduction in distress is only one aspect of evaluating the clinical impact of an antidepressant drug. Thus they end their report by speculating that the reason that serotonin reuptake blocking agents are more *effective* in generalized anxiety disorder and obsessive–compulsive disorder than drugs which do not have serotonin reuptake blocking activity is precisely *because* serotonin reuptake blocking agents induce frontal lobe dysfunction. Since efficacy is not distinguished from toxicity in this final comment, it looks like the authors cannot or do not wish to make such a distinction.

How to Distinguish Between Toxic Psychological Drug Effects and Therapeutic Psychological Drug Effects: Neglected Methodological Innovations

Discerning the psychological effects of a psychoactive drug is rendered even more complicated when the drug is being used as medicine for "psychiatric conditions." One direct way to shed light on the question "therapeutic drug effect or toxic/adverse drug effect?" is to administer what are thought of as therapeutic doses and durations of a psychoactive drug to normal (no diagnosis) subjects. At issue in such an investigation is whether the drug brings about untoward psychological alterations which could be construed as symptom relief if the drug was used for clinical purposes (e.g., if the drug induces emotional indifference in non-patients, *the same effect* could be seen by researchers as salutary/therapeutic when it occurs in people who come for treatment complaining of feeling distressed, upset, despairing, agitated, blue, despondent, etc.).

The foregoing suggestion is drawn from a research report concerning the effects of lithium carbonate on normal subjects published in 1977 by Judd, Hubbard, Janowsky, Huey, and Atwell (the lead author, Lewis L. Judd, became Director of NIMH for a three year period in the late 1980s). Judd et

[10]The authors describe all the patients selected for this adverse fluoxetine reaction report as displaying the following frontal lobe injury symptoms: apathy, flatness of affect, lack of emotional concern, childishness and euphoria, socially inappropriate behavior, and difficulty in foreseeing the outcome of an action (p. 345).

al.'s report has a complicated structure as a text in that the non-experimental findings presented are actually more interesting and to the point than the experimental data. The non-experimental observations are presented as introductory remarks and as the impetus for the study. Two sources of non-experimental observations are presented. The first source is presented following the comment that few studies have investigated the effects of therapeutic doses and durations of lithium carbonate on normal subjects. The point here is that it may be highly misleading to study dosages and durations of drug administration which are smaller and briefer than those used in clinical treatment. Judd et al. make this point by citing a 1968 report concerning lithium carbonate administration conducted by Schou. In this study sub-clinical doses of lithium carbonate over a seven day period produced few detectable effects in normal subjects. However, Schou and his research associates took approximately 1,850 mg/day themselves for several weeks and reported a variety of adverse effects (summarized in Judd et al.): increased mental effort in initiating physical tasks (inertia), indifference, passivity, decreased response to environmental stimuli, being separated from environmental stimuli by a "glass wall," etc. It is impossible to read this list without immediately thinking of how such effects would impress clinicians prescribing lithium carbonate to treat mania, psychotic agitation, etc.

The second source of non-experimental observations derived from a prior study conducted by Judd and Hubbard (1975) in which nine normal subjects were maintained on therapeutic doses of lithium carbonate for a two week period in a non-blind format. During the course of this two-week period Judd and Hubbard were impressed by the dulling and blunting effects of lithium carbonate on their normal male subjects. This sort of observation concerning the effects of lithium carbonate on individuals who are not ill and are not being treated for anything also raises the suspicion that it is the behavioral toxicity (Summerfield, 1978) of so-called therapeutic doses of lithium carbonate which bring about what clinicians and researchers construe as its medicinal effects.

Judd et al. conclude their introduction by stating their intention to determine if the nonblind, "anecdotal" observations they cited could be "objectively demonstrated in a well-controlled clinical study." To accomplish this they carried out a double blind, placebo vs. lithium carbonate crossover study counterbalanced for order. The psychopharmacological data derived (mainly, see below) from a variety of self-rated inventories administered at the end of the two week placebo period and the two week lithium carbonate maintenance period. Compared to how subjects rated the inventories at the end of the placebo period, numerous significantly different mean differences (item by item) were obtained at the end of the lithium carbonate maintenance period, all in the direction of greater distress, dysphoria, disability, etc. Judd

et al. comment that the results of the foregoing well-controlled clinical study are "basically in agreement" with how Schou and his colleagues depicted the effects of lithium carbonate on themselves in their 1968 report. The difference is worth noting. While the long list of adjectives supplied to subjects by Judd et al. indicates (as they say) that lithium carbonate produces dysphoric and disabling changes in normals, the self-selected descriptions by Schou and his colleagues indicate far more clearly how the effects of lithium carbonate could be seen by clinicians as *therapeutic* when administered to psychiatric patients who display certain symptoms. For example — keeping the treatment of mania in mind — the list of adjectives supplied by Judd et al. permits subjects to indicate that they feel more helpless, confused, and exhausted on lithium carbonate, and also less clearheaded; and feel that ideas do not flow as easily (these effects hardly seem beneficial under any circumstances), while the self-selected descriptions of Schou and colleagues depict lithium carbonate as bringing about inertia, indifference, passivity, decreased response to environmental stimuli, and feeling separated from environmental stimuli by "a glass wall." Thus while the experimental design permitted direct (vs. implicit) comparison between the drugged condition and the non-drugged condition, and also presumably controlled for subject expectations concerning drug effects (*via* the double-blind arrangement), the experimental design also controlled what subjects *could* say about the effects of lithium carbonate. Indeed, if Schou and his colleagues were *subjects* in Judd et al.'s study, then the anecdotal (and enlightening) information they conveyed about the effects of lithium carbonate on normal subjects would have been lost.[11]

An additional experimental manipulation incorporated into Judd et al.'s design allowed for light to be shed on what can be expected of researchers who do not know the subjects as individuals and whose observational opportunities during the course of the study are highly restricted when it comes to noticing drug effects. Before the study began each subject was asked to designate a "significant other" who would be called upon to render judgments about his (the subjects were all male) psychological condition at the end of each two week period (placebo vs. lithium carbonate). The significant others were of course also "blinded." The judgments made by each significant other were based upon whatever relationship and contacts the significant other had with each subject in their everyday lives. The judgments of the significant others were compared to judgments made by "trained observers" (presumably research colleagues) on the basis of "a short personal interaction" with subjects followed by simply watching them fill out the research instru-

[11]The lesson that can be taken here is that the ideal of obtaining a total picture of the subjective changes brought about by a psychoactive drug must include an opportunity for subjects to *speak for themselves*. Judd et al. explain that they used so many self-rated inventories precisely for the purpose of pursuing a total picture of lithium carbonate-induced subjective changes.

ments. The task for the significant others and the trained observers consisted of rating a list of adjectives describing an individual's present psychological state thought to be amenable to behavioral observation, e.g., happy, angry, grouchy, drowsy. Each subject rated the same list of adjectives for themselves on the same day the observers made their evaluations. Although subjects indicated by their self-evaluations that the effects of lithium carbonate were profoundly noticeable and dysphoric, the trained observers were unable to distinguish differences between the subjects' behavior or mood on or off lithium carbonate.[12] By contrast, the significant other evaluations were highly consistent with the subjects' self-ratings on and off lithium carbonate.

However, it should not be overlooked that this finding — as suggestive as it is for what can be expected of researcher observations — does not address the problem of recognizing drug effects which are not reported to the researcher by the subject. The comparison just described between trained observer evaluations and significant others evaluations depended upon subject self-evaluations as a standard. No observer evaluations were requested which bear on the issue of the possible disparity between what an observer can notice and what a drugged person can notice about himself or herself. The prior observations of Judd and Hubbard were of this nature, that is, observations *about* lithium carbonate-drugged subjects which would probably not be forthcoming from the subjects themselves — e.g., "we anecdotally noted an overall dulling and blunting of various personality functions. . . " (p. 347). The problem of recognizing subtle drug effects is critical for developing a comprehensive and realistic picture of just what is brought about by psychoactive drugs used as medicine. It is important to notice that the problem is left dangling in Judd et al.'s study.

A Further Methodological Innovation: Drug-Free, Retrospective Descriptions

In the Judd et al. study reviewed above no information about drug effects was solicited from subjects following discontinuation of the drug after the subjects had had an opportunity to recover from its effects. This is consistent with standard procedure in RCTs, which as far as I can tell never solicit information from subjects about drug effects following discontinuation. Post-drug recovery descriptions of drug effects, although admittedly complicated by the necessity to draw upon memory, may add information which is not available from any other source. Judd et al. do not indicate when Schou and his

[12]The apparent inability of the research confederates here to discriminate between "on lithium carbonate" and "on placebo" should not be construed as support for the effectiveness of researcher "blinding" in RCTs. In the latter (treatment) situation, researchers directly ask subjects about the emergence of adverse drug effects on a weekly basis. The research observers in the Judd et al. study were not able to ask subjects about drug effects.

research associates created the written descriptions of their experience of lithium carbonate which appeared in Schou's 1968 report. The descriptions Schou provided make it reasonable to wonder whether they were written after the effects of lithium carbonate wore off (increased mental effort in initiating physical tasks, indifference, passivity, etc.). The point is that some drug effects may interfere with realistically providing witness until the person is no longer in the grip of the drug. For example, a group of British researchers (Golombok, Pamala, and Lader, 1988) report that patients on long-term benzodiazepine treatment often comment that they did not realize how psychologically impaired they had become until they had successfully withdrawn.

In a recent critical review paper concerning neuroleptic ("antipsychotic") drugs, Cohen (1997) provides a vivid illustration of the importance of retrospective drug depictions which derives from what is apparently the first use of chlorpromazine in psychiatry. The first use occurred in 1951 when French psychiatrist Leon Chertok injected an unspecified amount of chlorpromazine into a colleague (Cornelia Quarti) and voice-recorded her comments. It is evident from Quarti's written account that she used the voice-recording days later as an aid in reconstructing what she experienced, and it is equally evident from her written account that the consequences of chlorpromazine rendered her unable to effectively produce a report while in the grip of the drug. My own (unpublished) interviews with people who have in the past been treated with antidepressants or anxiolytics also strongly indicate how informative it can be to obtain accounts of drug effects once the person is no longer under the influence of the drug. For example, during the last of a series of five interviews with a man who had been treated for depression with Prozac, he remarked that he could not have talked with me about his life with such feeling and in such depth when he was on Prozac.[13] His retrospective view was that Prozac rendered him too emotionally numb, detached, disinterested, to become engaged in such dialogue with anyone. He thinks he was at times somewhat cognizant of the psychological alterations which Prozac brought about, but his overriding interest when he began taking Prozac was in obtaining relief from the intensity of his feelings, and in this sense his view is that Prozac brought him that relief. It is not clear what he would have said about his drug-induced state while on Prozac. It is easy to imagine that the relief he experienced from his feelings when he began on Prozac — although retrospectively depicted primarily in terms of numbness,

[13]I assume his official diagnosis when he was prescribed Prozac would have been major depressive disorder. But to my mind the term depression and the criteria for its diagnosis in any of the recent DSMs fail to indicate the complexity of his distress as he described it to me. Certainly his mood was "depressed" — he was hardly in high spirits — but this does not begin to capture the multifaceted nature of his anguish when he desperately sought out a psychiatrist and pleaded for medication.

detachment, etc. — could have been seen in an RCT as reduced depressive symptomatology. This takes us full circle to the Hoehn–Saric et al. (1990) report in which the question of how to interpret a drug-produced reduction in distress was posed. Although the individual I interviewed had to rely on memory to tell me about his experience on Prozac, his retrospective account does not appear to be a description of a specific "antidepressant" effect. It is not necessary to decide which account of drug effects (contemporaneous or retrospective) is more or less trustworthy, valid, free of bias, etc. As Zinberg (1976) has pointed out in the context of studying drug-induced alterations in consciousness, every methodological choice influences the quality and quantity of information obtained. The aim of systematic study should be to acquire information in a variety of ways so that the complexity of the subject can emerge and consistencies and inconsistencies can be noted. If retrospective drug depictions appear inconsistent with self-reports while the drug is biologically active in the person, this feature of psychopharmacology should not be obscured.

To sum up this section, it would appear that two conclusions are warranted: (i) the "side effects" established in psychopharmacotherapy RCTs do not by any means tell the whole story concerning even blatant adverse drug reactions, (ii) the design of psychopharmacotherapy RCTs essentially rules out the detection of subtle adverse psychological effects by default (non-investigation), thereby rendering the meaning of drug-induced symptom reduction highly ambiguous.[14]

The Non-Control of Bias

It is now de rigueur for psychopharmacotherapy researchers to report that they took steps to blind the treating psychiatrists as well as subjects as to who was on what substance. The point is to keep both parties in ignorance as to the real identity of what was used for treatment purposes (the investigational

[14]After completing the present manuscript, I came across Jonathan Cole's 1960 book chapter entitled "Behavioral Toxicity." Since I am making the same points in 1999, it hardly needs to be said that psychopharmacotherapy research has for the most part ignored the substantive issues Cole raised in 1960. Specifically: (i) "The most impressive characteristic of the psychiatric drug literature is the absence of serious concern about adverse effects these drugs may be having upon behavior" (p. 170); ". . . when discussing side effects, concern [is directed] primarily with non-behavioral forms of toxicity . . ." (p. 171); (ii) ". . . clinicians seeing outpatients may see them only in the interview situation and may not be able to detect adverse behavioral phenomena which occur only when the patient is at home or at work" (p. 171); ". . . studies which take into account both the patient's subjective response to [drug] therapy and information from other informants concerning alterations in his behavior are obviously needed" (p. 174); (iii) ". . . almost all published studies on normal subjects deal with single doses of the drug in question and may have only limited relevance to the clinical situations where chronic drug administration is the rule" (p. 179).

drug or the placebo) and thereby to rule out bias. The most obvious way to do this is to make the placebo (pill, capsule, injection) look just like the active medication, so that no one can tell just by looking who is getting what. This seems necessary and obvious as far as it goes, but one has to suspect that a vast array of unwanted drug effects (e.g., parkinsonian rigidity and tremor) are not likely to keep either the treating psychiatrist or the subject in the dark as to who is actually taking the active medication versus who is taking placebo. The arrangement of the study as a whole can in fact be seen to virtually guarantee that the formal blinding procedure will be undone during the course of the investigation.

I will emphasize for the present how the treating psychiatrist becomes "unblinded" during the course of research. Since most investigations rely exclusively upon the treating psychiatrists' assessments of subjects' clinical status for "the data" (vs. subjects' self-evaluations), it is first and foremost necessary to show how the investigation permits the treating psychiatrists to see through the double blind. The flaw in the ordinary double bind arrangement is that treating psychiatrists meet weekly with subjects in order to ask about symptoms and side effects. This means that the treating psychiatrists will hear complaints concerning drug effects that they know are expected consequences of the active medication being used in the investigation, and they can directly observe what effect dosage adjustment has on these complaints during the course of the investigation (impotence, nausea, intense headaches, tremor, etc.). It is precisely this arrangement which permits the treating psychiatrists to see through the formal double blind.

Since the point of the standard procedure for blinding treating psychiatrists and subjects is to prevent bias from affecting results, the question must be asked why further — and obvious — steps are not taken to generate clinical data that are free or at least freer of the distorting effects of bias. I emphasize that further obvious steps in this direction have been suggested within the psychiatric literature from the 1950s onward, that is, since the introduction of chlorpromazine into psychiatry made it apparent that blinding anybody involved in a clinical trial was going to present a formidable challenge. The following steps could be taken to further protect clinical data from being infiltrated by bias. Neither cost nor inconvenience are persuasive counterarguments for refraining from taking measures to guard against bias, since the safety and effectiveness of any medication can only be established on the basis of controlled clinical studies that are scientifically adequate (i.e., do not produce spurious results). Additional measures are as follows:[15]

[15]I have drawn upon a variety of sources here: Clark, 1956; Dinnerstein, Lowenthal, and Blitz, 1966; Gilmour, 1977; Greenberg, Bornstein, Greenberg, and Fisher, 1992; Lambert, Hatch, Kingston, and Edwards, 1986; Marholin and Phillips, 1976; Prien and Levine, 1984; Raskis and Smarr, 1957; Smith, Traganza, and Harrison, 1969; Summerfield, 1978; Thomson, 1982.

1. The observer-based clinical data should not be in the hands of the treating psychiatrist who is understandably biased in favor of the medication he/she is using, and who meets weekly with subjects to review medication-produced side effects. The observer-based clinical data should be in the hands of clinicians who are not treating the subjects, who do *not* systematically review the medication's side effects with the subjects, who do *not* adjust dosages for the subjects, but who meet periodically with subjects to assess their clinical status.

2. The duration of the clinical trial should represent how the medication is actually used in clinical practice — the duration of the trial should *not* represent what is in clinical practice only the very beginning phase of treatment with the medication. In fact this single modification — extension of the trial to conform to how the drug is used in clinical practice — would obviate a number of severe threats to validity which emerge from the typically very brief psychopharmacotherapy study, notably (a) the total psychological effects on the subject which derive from participating in an experimental investigation of a promising medication — but which are all essentially irrelevant to ordinary clinical practice — would have an opportunity to decay over an adequate period of time, (b) the active medication's amplified placebo effect — which is typically entirely uncontrolled in virtue of the absence of an *active* placebo treatment group — would likewise have an opportunity to decay over an adequate period of time, (c) the true incidence of adverse drug effects during a time period which conforms to actual clinical practice would have an opportunity to emerge, (d) a realistic picture of the extent to which tolerance for the medication's desired clinical effects develops over relevant time periods would have an opportunity to emerge, (e) the extent to which the drug's total adverse effects precipitate patient discontinuance over relevant time periods would have an opportunity to emerge.

3. The subject's own view of the positive and negative effects of the investigational drug should find direct expression in the clinical data *without* being filtered through the researchers. For example, the most commonly used research instrument for evaluating clinical efficacy of antidepressant medications is the Hamilton Psychiatric Rating Scale for Depression (HAM-D), which according to Hamilton (1960) can be used to quantify the results of an interview with a patient who has *already* been diagnosed as suffering from a depressive type affective disorder. Hamilton does not specify just how the original diagnosis should be made, but his remarks throughout indicate this is essentially a matter of clinical judgment. He does explicitly recognize that many of the items (symptoms) on the 21 item inventory are not specific to depression. Indeed, arguably only one item, depressed mood, is specific to depression. For example, it is patently obvious that sleeping difficulties, anxiety, agitation, sundry gastrointestinal symptoms, paranoid symptoms, etc. are

not specific to depression. Hamilton evidently believed in the existence of discrete psychiatric illnesses, but he nonetheless recognized that in practice patients given one diagnosis concurrently display diverse symptoms of other diagnoses. Presumably this is why he made the point that symptom-rating scales should not be used for the purpose of diagnosis, i.e., "Thus the schizophrenic patients should have a high score on schizophrenia and comparatively small scores on other syndromes. In practice, this does not occur" (p. 56). Despite his apparent belief in the reality of discrete diagnostic entities, it could be argued that his inclusion of so many items that are not specific to depression (as many as 20 out of 21) on his "rating scale for depression" reveals that his theoretical commitment to discrete psychiatric illnesses was incompatible with his desire to create a tool which had practical clinical value. The implications of this should be apparent by now.

It has been pointed out repeatedly by psychiatric researchers (e.g., NIMH researchers Prien and Levine [1984]) that the HAM–D, precisely in virtue of the large number of items which appear on it which are not specific to depression, should *not* be used to assess the antidepressant activity per se of a medication.[16] Nevertheless, change scores on the HAM–D remains the standard in psychiatric drug research for assessing the antidepressant effect of a medication. Use of the HAM–D may result in positive findings as far as the researchers are concerned without substantially influencing feelings and/or functional capacities of paramount interest to subjects. In contrast to the HAM–D, the Beck Depression Inventory (BDI) is filled out directly by experimental subjects themselves, and appears to address "mood and feeling life" to a far greater extent than the HAM–D (Lambert, Hatch, Kingston, and Edwards, 1986). I do not mean to minimize the importance of (unbiased) observer based clinical data. The main point is that discrepancies between investigators' views of what the medication accomplished and subjects' own views should not be hidden by the investigators' control of how drug efficacy is assessed.

---

[16] In a chapter concerning the assessment and measurement of depression published by the American Psychiatric Association Press, Wetzler and van Pragg (1989) point out that it is inappropriate to use the HAM–D to indicate the antidepressant action of a drug because "it is sensitive to many facets of psychopathology that are not specifically depressive" (p. 80). In other words, the usefulness of the HAM–D as Hamilton (1960) envisioned it was to "quantify the results of an interview" with a patient *already* diagnosed as suffering from an affective disorder on the basis of clinical judgment, *not* (as he specifically said) to identify depression or to treat the scale *as if* it measured a univocal dimension of depression. Thus Hamilton did *not* present his scale as having either construct or discriminant validity regarding depression per se. In light of the HAM–D items devoted to insomnia, anxiety, agitation, and diurnal variation, it is more or less evident that drugs which produce sedation and tranquilization will appear to have an antidepressant effect if the HAM–D is mistakenly used to measure a univocal dimension of depression.

The difference between what is found when further steps are taken to protect against bias and what is found based upon the typical or standard protocol is strikingly illustrated in the 1989 NIMH investigation of short term treatments for unipolar depression (Elkin et al., 1989). The unprecedented advantage of this study lies in the fact that clinical data regarding the effectiveness of antidepressant medication (imipramine) obtained in the usual manner can be directly compared to data obtained under more stringent protections. The "more stringent protections" consisted of (a) obtaining HAM–D scores from independent clinical evaluators (in addition to HAM–D scores from the treating psychiatrists), (b) extending the duration of treatment to 16 weeks and (c) the BDI (scored directly by the subject him/herself) was used as a clinical measure as well as the HAM–D (scored by a clinician on the basis of interviewing the subject).

The pertinent findings as far as the present discussion is concerned are as follows.

1. Based on HAM–D scores by independent clinical evaluators (not treating psychiatrists), the imipramine treated group did no better than the placebo treated group or two other groups treated with either cognitive behavioral therapy or interpersonal therapy.

2. By contrast, when HAM–D scores generated by the treating psychiatrists for the imipramine treated group and the placebo treated group were compared (the same psychiatrists treated both groups), the imipramine treated group did significantly better than the placebo treated group.

3. On the BDI, the placebo group was not significantly different at the conclusion of treatment than the imipramine treated group.

4. Based upon the standard of recovery on the HAM–D at the conclusion of treatment which consisted of a score ≤ 6, the treating psychiatrists rated 78% of imipramine treated subjects who received at least 3.5 weeks of treatment as recovered. This contrasts to 47% of the same group evaluated as recovered by the independent clinical evaluators (Table 2, p. 976). Elkin et al. do not provide a test of statistical difference between these two figures, but they do comment that the recovery rate as judged by the treating psychiatrists was striking.

5. Based upon Elkin et al.'s commentary, it appears that they not only took the unusual step of asking the treating psychiatrists to state their opinion as to who was on imipramine and who was on the inert placebo, but went even further to report what they learned by taking this step. Unfortunately, they provide no pertinent figures on this matter, but the text makes it clear that the treating psychiatrists could discern who was treated with imipramine and who was treated with placebo. Thus psychiatrists could see through their own "blinded" condition and, as a matter of fact, rated the imipramine treated subjects as substantially more improved than did the independent

clinical evaluators. The importance of these two disparities (degree of blindness and ratings on the HAM–D) between the independent clinical evaluators and the treating psychiatrists is precisely the difference between finding no advantage for imipramine over placebo and a "very favorable" finding for imipramine with "usual psychopharmacology standards" (p. 978).

On the face of it — given the clear disparity in the Elkin et al. 1989 NIMH study between what is found using the "usual psychopharmacology standards" and what is found when steps are taken to further protect against bias — one would have expected an immediate and pervasive change in the conduct of psychopharmacotherapy research. Obviously this has not happened. The problem of course is that investigator bias begins to operate immediately, as soon as a psychopharmacotherapy investigation is planned by a group with a vested interest in a particular drug (for a general discussion of this issue in medicine, see Porter, 1992). The NIMH itself of course could refuse to fund psychopharmacotherapy studies which do not incorporate appropriate methodological modifications, but this also has not occurred.

Evidently Elkin et al. were not at all pleased with their own data, which showed that imipramine was not more effective than inert placebo. In what looks like an attempt to salvage something for imipramine (by extension, drug treatment), the investigators conducted what they called a secondary analysis of their outcome data in which the entire sample of subjects was divided into more and less severely disturbed subgroups based upon their pretreatment scores on the HAM–D and the Global Assessment Scale (GAS). The end result of the secondary analysis is that Elkin et al. were able to report some advantages for imipramine treatment in the more severely disturbed group. Having re-analyzed their outcome data in the foregoing manner, Elkin et al. go on to discuss the results of the secondary data analysis as a finding of the study later in the paper. Since this study has become widely known, it is important to point out why the finding regarding severity of disturbance and imipramine treatment is not a bona fide outcome of the investigation. In the first place, the secondary analysis is just that — the product of data analysis once the outcome data were already "in." Also, Elkin et al. appear to have forgotten that potential subjects were *excluded* from the study if they displayed "a clinical state inconsistent with participating in the research protocol, e.g., current active suicide potential or need for immediate treatment" (p. 972). This exclusion criterion obviously selected out highly disturbed subjects.

To What Should Drug Treatment be Compared?

Random assignment of subjects to medication treatment or placebo treatment creates problems which are inadequately addressed in conventional psychopharmacotherapy research. As Gram (1994) points out in his review

paper concerning fluoxetine, most subjects for psychopharmacotherapy studies in the United States are recruited on the basis of public advertisements which offer treatment with a promising new drug. But the necessity to obtain informed consent from subjects obliges researchers to explain that (based on random assignment) treatment may actually consist of placebo administration. It is hardly surprising then that subjects who are assigned to the placebo treatment group typically drop out of the study at a higher rate than subjects in the active medication group, an outcome which necessitates a blatantly inappropriate form of data analysis (discussed below).

If the purpose of the placebo group is to keep treating psychiatrists in ignorance as to whom is on what and therefore to rule out experimenter bias, I believe I have already shown that placebo is inadequate for this purpose. It is not likely that random assignment to medication or placebo keeps subjects ignorant either, since (a) dramatic adverse conditions will soon arise in a proportion of subjects following the initiation of treatment, and (b) experienced psychiatric drug users are accepted into research studies (e.g., at least 55% of the subjects in the Elkin et al. 1989 NIMH study). In terms of bias control, then, the placebo treatment group appears to serve no purpose.

If the purpose of the placebo treatment group is to show that the medication is superior in terms of relieving suffering, two objections can be raised. The first objection is that since the placebo fails (compared to the active medication) to keep subjects in the study for the (brief) planned period of treatment, the active medication de facto is *not* being evaluated in terms of its advantage over placebo. As Gram (1994) notes, the efficacy of a psychiatric medication is usually evaluated on the basis of what is known as end point analysis, that is, the last available clinical evaluation for each subject is taken to indicate treatment outcome. Drawing upon the studies and figures Gram provides, in six-week studies of fluoxetine 61% of the placebo group and 45% of the fluoxetine group dropped out of treatment (Table 2, p. 1357). Under "Methodologic Problems," Gram notes that analysis of the clinical outcome data depends upon dropout pattern, but he does not observe that a greater dropout rate for the placebo group than the fluoxetine group means that end point analysis *excludes* much of the placebo effect from evaluation. The upshot is that end point analysis does not actually compare the clinical effects of the medication with the clinical effects of the placebo. Gram only presents efficacy results from end point data (Table 1, p. 1357); no data are reported based on subjects who completed the planned time period of treatment in each experimental condition (this comparison is generally much less favorable for the investigational drug vs. placebo, e.g., Cohn and Wilcox's [1985] investigation of fluoxetine treatment for major depressive disorder).

The second objection is that showing the superiority of medication over placebo (or even another medication) is *not* sufficient for the purpose of jus-

tifying the clinical use of a drug as a treatment. The decision to actually use a drug as a first resort treatment must be based upon a convincing demonstration that the drug is superior to available psychological forms of treatment. The reason for this should be obvious: all psychiatric medications (psychoactive substances which alter neurophysiology) are dangerous to at least some people who take them, and it is impossible to predict in advance who is especially vulnerable to the drug's medical and behavioral toxic effects. As obvious as this point seems it is nevertheless the case that ignoring the existence of psychological forms of treatment is one of the bedrock traditions in psychopharmacotherapy research. In a 1969 publication Smith, Traganza, and Harrison reviewed more than 2000 articles concerning the efficacy of antidepressant medications published during the period 1955–1966. Of the original 2000 articles 473 contained sufficient statistical information for inclusion in the final review. Of these 473 studies Smith et al. found a total of two studies which compared antidepressant medication to psychotherapy. Of these two none found antidepressant medication to be more effective than psychotherapy (Table 23, p. 9).

The combination of recruiting people for drug treatment and then thwarting their desire to receive treatment (by assignment to placebo treatment), inadequately blinding the treating psychiatrists who provide the drug-efficacy relevant data, and avoiding comparing drug treatment to psychological treatment is the situation which the *investigators* contrive. The random assignment of subjects to placebo treatment *sounds* scientific, but no step is intrinsically "scientific" regardless of context. It is hardly scientific to contrive a comparison *treatment* condition which will not keep the majority of subjects in the condition for even a planned six-week course of treatment, meanwhile evading the form of treatment (i.e., psychotherapy) to which the drug should be compared. It would be far more sensible to not blind clinicians or subjects. If the treatment period was substantially extended, if psychotherapy was provided as a comparative form of treatment, and if the clinical evaluators (who are not the treating psychiatrists) were truly independent, disinterested parties, then the treatment outcome data would be far more clinically meaningful. As matters stand, it is not in the best interests of patients to rely on studies in which drug effectiveness has the tortured meaning it does in virtue of how the studies are conducted.[17]

[17]Despite all the researcher contrived advantages for the investigational drug, Gram notes that failure to publish the results of (negative) clinical trials is a general problem in the area of antidepressant medication research. He specifically cites unpublished clinical data concerning fluoxetine sponsored by Eli Lilly, although the reader must know how to interpret the citation he provides, i.e., the reader must already know that the lead author, C.M. Beasley, is an employee of Eli Lilly and that the article attempts to refute concerns about fluoxetine-induced suicide by presenting previously unpublished information.

Conclusion

Clinicians, policy makers, and others who believe that a mountain of scientific evidence exists which shows that distressed people benefit from psychopharmacotherapy are in the unenviable position of endorsing "usual psychopharmacology standards" (Elkin et al., 1989, p. 978). Since the introduction of chlorpromazine into North American psychiatry in the early 1950s a steady but thin stream of publications within psychiatry journals has called attention to the inescapable fact that experimental results in psychopharmacotherapy research, no matter how congenial to psychiatry, cannot be divorced from the unsound experimental methods upon which they depend. Calls for substantive reform have been in vain. In 1969 the Smith, Traganza, and Harrison review of antidepressant clinical research, which was sponsored by the Psychopharmacology Branch of the NIMH, concluded that "the methodology of drug research is of more significance to the outcome of a clinical trial than is the drug being studied" (p. 19). Twenty years later Elkin et al., another NIMH sponsored research team, repeated that sentiment, namely that the usual methods employed in psychopharmacotherapy research create a picture of drug treatment efficacy which is not sustainable under more adequate experimental conditions. It cannot be overlooked that an empirical demonstration of the extent to which the antidepressant drug held up as the standard of effectiveness in the industry (imipramine) depends upon unsound experimental methods has failed to reform accepted research practices. My own review of psychopharmacotherapy research shows, I believe, that the problems facing researchers are far more complex and have been met far less adequately than is recognized in psychiatric publications. The overall conclusion which can be drawn from the present review is that the (by now) tens of thousands of psychopharmacotherapy studies which have been conducted *in the usual manner* provide no reason to think that much is known about the usefulness of treating distressed people pharmacologically, even symptomatically or in the short run. The problem of distinguishing between genuinely therapeutic effects of a psychoactive substance and subtle adverse/impairing effects is probably the most difficult challenge to clinical research. As discussed, this problem is for all intents and purposes not even recognized as such in controlled psychopharmacotherapy research. I realize that some readers may become stuck on the (counter-intuitive) proposition that drug-induced reports of *lessened* distress may be a cause for alarm rather than a mark of progress in medical psychiatry. I can only reiterate that the phenomenon of neuropathologically-induced unconcern and related psychological phenomena is well known in clinical neurology (a vivid example is provided by Sachs [1987], Chapter 13). Psychopharmacotherapy researchers have not taken the possibility of drug-

induced impairment in realistic self-monitoring/self-appraisal seriously, perhaps because research would become substantially more complicated and costly if it was. But of course ignorance of the true state of affairs will not save patients from iatrogenic damage. The history of neuroleptic-induced tardive dyskinesia makes this point only too well.

References

Ananth, J. (1984). Physical illness and psychiatric disorders. *Comprehensive Psychiatry, 25*, 586–593.

Andreason, N.D., and Carpenter, Jr., W.T. (1993). Diagnosis and classification of schizophrenia. *Schizophrenia Bulletin, 19*, 25–40.

Andrews, G., Steward, G., Morris–Yates, A., Holt, P., and Henderson, S. (1990). Evidence for a general neurotic syndrome. *British Journal of Psychiatry, 157*, 6–12.

Antonuccio, D.O., Danton, W.G., and DeNelsky, G.Y. (1995). Psychotherapy vs. medication for depression: Challenging the conventional wisdom with data. *Professional Psychology: Research and Practice, 26*, 574–585.

Austin, S.C., Stolley, P.D., and Lasky, T. (1992). The history of malariotherapy for neurosyphillis. *Journal of the American Medical Association, 268*, 516–519.

Bleuler, M. (1978). *The schizophrenic disorders: Long-term patient and family studies*. New Haven, Connecticut: Yale University Press.

Bleuler, M. (1991). The concept of schizophrenia in Europe during the past one hundred years. In W.F. Flack, Jr. and D.R. Miller (Eds.), *What is schizophrenia?* (pp. 1–16). New York: Springer–Verlag.

Bogerts, B. (1993). Recent advances in the neuropathology of schizophrenia. *Schizophrenia Bulletin, 19*, 257–271.

Bolos, A.M., Dean, M., Lucas–Derse, S., Ramsburg, M., Brown, G.L., and Goldman, D. (1990). Population and pedigree studies reveal a lack of association between the dopamine D_2 receptor gene and alcoholism. *Journal of the American Medical Association, 26*, 3156–3160.

Breggin, P.R. (1983). *Psychiatric drugs: Hazards to the brain*. New York: Springer Publishing Company.

Breggin, P.R. (1991). *Toxic psychiatry*. New York: St. Martin's press.

Breggin, P.R. (1993). Parallels between neuroleptic effects and lethargic encephalitis. *Brain and Cognition, 23*, 8–27.

Breggin, P.R. (1997). *Brain disabling treatments in psychiatry*. New York: Springer Publishing Company.

Byerley, W.F., Reinherr, F.W., Wood, D.R., and Grosser, B.I. (1988). Fluoxetine, a selective serotonin uptake inhibitor, for the treatment of outpatients with major depression. *Journal of Clinical Psychopharmacology, 8*, 112–115.

Caine, E.D., and Shoulson, I. (1983). Psychiatric syndromes in Huntington's Disease. *American Journal of Psychiatry, 140*, 728–733.

Carpenter, Jr., W.T., McGlashan, T.H., and Strauss, J.S. (1977). The treatment of acute schizophrenia without drugs: An investigation of some current assumptions. *American Journal of Psychiatry, 134*, 14–20.

Carson, R.C., and Sanislow, A. (1993). The schizophrenias. In P.B. Sutker and H.E. Adams (Eds), *Comprehensive handbook of psychopathology* (pp. 295–333). New York: Plenum Books.

Ciompi, L. (1997). The Soteria–concept. Theoretical bases and practical 13-year-experience with a mileu–therapeutic approach of acute schizophrenia. *Psychiatrica et Neurologia Japonica, 99*, 634–650.

Clark, L.D. (1956). Evaluation of the therapeutic effects of drugs in psychiatric patients. *Diseases of the Nervous System, 17*, 282–286.

Cochrane, A.L. (1972). *Effectiveness and efficiency*. London: Burgess & Son Ltd.

Cohen, D. (1997). A critique of the use of neuroleptic drugs in psychiatry. In S. Fisher and R.P. Greenberg (Eds.), *From placebo to panacea: Putting psychiatric drugs to the test* (pp. 173–228). New York: John Wiley and Sons, Inc.

Cohn, J.B., and Wilcox, C. (1985). A comparison of fluoxetine, imipramine, and placebo in patients with major depressive disorder. *Journal of Clinical Psychiatry, 46,* 26–31.

Cole, J.O. (1960). Behavioral toxicity. In L. Uhr and J.G. Miller (Eds.), *Drugs and behavior* (pp. 166–183). New York: John Wiley and Sons, Inc.

Dinnerstein, A.J., Lowenthal, J., and Blitz, B. (1966). The interaction of drugs with placebos in the control of pain and anxiety. *Perspectives in Biology and Medicine, 10,* 103–117.

Dubos, R. (1965). *Man adapting.* New Haven: Yale University Press.

Dumont, M.P. (1984). The nonspecificity of mental illness. *American Journal of Orthopsychiatry, 54,* 326–334.

Duster, T. (1984). A social frame for biological knowledge. In T. Duster and K. Garrett (Eds.), *Cultural perspectives on biological knowledge* (pp. 1–40). Norwood, New Jersey: Ablex Publishing Co.

Duster, T. (1990). *Backdoor to eugenics.* New York: Routledge.

Easton, K., and Link, I. (1986). Do neuroleptics prevent relapse? Clinical observations in a psychosocial rehabilitation program. *Psychiatry Quarterly, 58,* 42–50.

Edwards, J.H. (1991). The formal problems of linkage. In P. McGuffing and R. Murray (Eds.), *The new genetics of mental illness* (pp. 58–69). London: Reed International Books.

Eisenberg, L. (1980). What makes persons "patients" and patients "well"? *The American Journal of Medicine, 69,* 277–286.

Elkin, I., Shea, M.T., Watkins, J.T., Imber, S.D., Sotsky, S.M., Collins, J.F., Glass, D.R., Dilkonis, D.A., Leber, W.R., Docherty, J.P., Fiester, S.J., and Parloff, M.B. (1989). National Institute of Mental Health treatment of depression collaborative research program. *Archives of General Psychiatry, 46,* 971–982.

Engel, G.L. (1977). The need for a new medical model: A challenge to biomedicine. *Science, 196,* 129–136.

Fisher, C.M. (1989). Neurologic fragments II. Remarks on anosognosia, confabulation, memory, and other topics. *Neurology, 39,* 127–132.

Frances, A.J., First, M.B., Widiger, T.A., Miele, G.M., Tilly, S., Davis, W.W., and Pincus, H.A. (1991). An A to Z guide to DSM–IV conundrums. *Journal of Abnormal Psychology, 100,* 407–412.

Gevirtz, R.N., Hubbard, D.R., and Harpin, E.R. (1996). Psychophysiologic treatment of lower back pain. *Professional Psychology: Research and Practice, 27,* 561–566.

Gilmour, R. (1977). Psychological factors in clinical trials. In F.N. Johnson and S. Johnson (Eds). *Clinical trials* (pp. 147–161). London: Blackwell Scientific Publications.

Golombok, S., Parmala, M., and Lader, M. (1988). Cognitive impairment in long–term benzodiazepine users. *Psychological Medicine, 18,* 365–374.

Gram, L.F. (1994). Fluoxetine. *The New England Journal of Medicine, 331,* 1354–1361.

Greenberg, R.P., Bornstein, R.F., Greenberg, M.D., and Fisher, S. (1992). A meta–analysis of antidepressant outcome under "blinder" conditions. *Journal of Consulting and Clinical Psychology, 60,* 664–669.

Greenberg., R.P., and Fisher, S. (1989). Examining antidepressant effectiveness. In S. Fisher and R.P. Greenberg (Eds.), *The limits of biological treatment for psychological distress* (pp. 1–37). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry, 23,* 56–62.

Heinrichs, R.W. (1993). Schizophrenia and the brain: Conditions for a neuropsychology of madness. *American Psychologist, 48,* 221–232.

Hoehn–Saric, R., Lipsey, J.R., and McLeod, D.A. (1990). Apathy and indifference in patients on fluvoxamine and fluoxetine. *Journal of Clinical Psychopharmacology, 10,* 343–345.

Hoffman, R.S. (1982). Diagnostic errors in the evaluation of behavioral disorders. *Journal of the American Medical Association, 248,* 964–967.

Hyman, S.E. (1996). Introduction from the new NIMH Director. *Psychotherapy and Rehabilitation Research Bulletin, 5,* 1–2.

Hyman, S.E., and Nestler, E.J. (1996). Initiation and adaptation: A paradigm for understanding psychotropic drug action. *American Journal of Psychiatry, 153*, 151–162.

Janicak, P.G., Davis, J.M., Preskorn, S.H., and Ayd, Jr., F.J. (1993). *Principles and practices of psychopharmacotherapy*. Baltimore: Williams & Wilkins.

Judd, L.L., and Hubbard, B.R. (1975). The effects of lithium carbonate on sodium phenobarbital induced high in normal subjects. *Psychopharmacology Communications, 1*, 631–639.

Judd, L.L., Hubbard, B.R., Janowsky, D.S., Leighton, H.Y., and Attewell, P.A. (1977). The effect of lithium carbonate on affect, mood and personality of normal subjects. *Archives of General Psychiatry, 34*, 346–351.

Karon, B.P. (1989). Psychotherapy vs. medications for schizophrenia. In S. Fisher and R.P. Greenberg (Eds.). *The limits of biological treatment for psychological distress* (pp. 105–150). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Karp, D.A. (1993). Taking anti-depressant medications: Resistence, trial commitment, conversion, disenchantment. *Qualitative Sociology, 16*, 337–359.

Karp, D.A. (1996). *Speaking of sadness*. New York: Oxford University Press.

Kendell, R.E. (1988). What is a case? Food for thought for epidemiologists. *Archives of General Psychiatry, 45*, 374–376.

Kendell, R.E. (1991). Schizophrenia: A medical view of a medical concept. In W.F. Flack, Jr., D.R. Miller, and M. Wiener (Eds.), *What is schizophrenia?* (pp. 59–72). New York: Springer–Verlag.

Kendell, R.E., and Brockington, I.F. (1980). The identification of disease entities and the relationship between schizophrenia and the affective psychoses. *British Journal of Psychiatry, 137*, 324–331.

King, R.A., Riddle, M.A., Chappell, P.B., Hardin, M.T., Anderson, G.M., Lombroso, P., and Kahill, L. (1991). Emergence of self-destructive phenomena in children and adolescents during fluoxetine treatment. *Journal of the American Academy of Child and Adolescent Psychiatry, 30*, 179–186.

Kirch, D.G. (1993). Infection and autoimmunity as etiologic factors in schizophrenia. *Schizophrenia Bulletin, 19*, 181–196.

Klerman, G.L. (1978). The evolution of a scientific nosology. In J.D. Shershow (Ed.), *Schizophrenia: Science and practice* (pp. 99–119). Cambridge, Massachusetts: Harvard University Press.

Klerman, G.L. (1988). Overview of the cross-national collaborative panic study. *Archives of General Psychiatry, 45*, 407–412.

Lambert, M.J., Hatch, D.R., Kingston, M.D., and Edwards, B.C. (1986). Zung, Beck, and Hamilton rating scales as measures of treatment outcome: A meta-analytic comparison. *Journal of Consulting and Clinical Psychology, 54*, 54–59.

Lewontin, R.C., Rose, S., and Kamin, L.J. (1984). *Not in our genes*. New York: Pantheon Books.

Lidz, T., Blatt, S., and Cook, B. (1981). Critique of the Danish–American studies of the adopted-away offspring of schizophrenic parents. *American Journal of Psychiatry, 138*, 1063–1068.

Lieberman, J.A., and Koreen, A.R. (1993). Neurochemistry and neuroendocrinology of schizophrenia: A selective review. *Schizophrenia Bulletin, 19*, 197–256.

Lipinski, Jr., J.F., Mallaya, G., Zimmerman, P., and Pope, Jr., H.G. (1989). Fluoxetine-induced akathisia: Clinical and theoretical implications. *Journal of Clinical Psychiatry, 50*, 339–342.

Marholin, D., and Philips, D. (1976). Methodological issues in psychopharmacological research. *American Journal of Orthopsychiatry, 46*, 477–495.

Marks, I. Swinson, R.P., Basoglu, M. Noshirvani, H., O'Sullivan, G, Lelliott, P.T., Kirby, M., McNamee, G., Sengun, S., and Wickwire, K. (1993). Alprazolam and exposure alone and combined in panic disorder with agoraphobia. *British Journal of Psychiatry, 162*, 776–787.

Marshall, E. (1994). Manic depression: Highs and lows on the research roller coaster. *Science, 264*, 1693–1695.

Mavissakalian, M., and Michelson, L. (1986). Two-year follow-up of exposure and imipramine treatment of agoraphobia. *American Journal of Psychiatry, 143*, 1106–1112.

McHugh, P.R. (1991). Schizophrenia and the disease model. In W.F. Flack, Jr., D.R. Miller, and M. Weiner (Eds), *What is schizophrenia?* (pp. 73–80). New York: Springer–Verlag.

Menn, A.Z., and Mosher, L. (1982). Forward. In H.S. Wilson, *Deinstitutionalized residential care for the mentally disordered* (pp. xi–xv). New York: Grune and Stratton.

Mesulam, M.M. (1990). Schizophrenia and the brain. *The New England Journal of Medicine, 322,* 842–844.

Milrod, B. (1995). The continued usefulness of psychoanalysis in the treatment armamentarium for panic disorder. *Journal of the American Psychoanalytic Association, 43,* 151–162.

Milrod, B., Busch, F.N., Hollander, E., Aronson, A., and Siever, L. (1996). A 23-year-old woman with panic disorder treated with psychoanalytic psychotherapy. *American Journal of Psychiatry, 153,* 698–703.

Mosher, L., and Burti, L. (1989). *Community mental health.* New York: W.W. Norton and Co.

Pam, A. (1990). A critique of the scientific status of biological psychiatry. *Acta Psychiatrica Scandinavica, 82* (Supplement 362), 1–35.

Polymeropoulos, M.H., Higgins, J.J., Globe, L.I., Johnson, W.G., Ide, S.E., DiLori, G., Sanges, G., Stenroos, E.S., Pho, L.T., Schaffer, A.A., Lazzarinin, A.M., Nussbaum, R.L., and Duvoisin, R.C. (1996). Mapping of a gene for Parkinson's Disease to chromosome 4q21–q23. *Science, 274,* 1197–1199.

Porter, R.J. (1992). Conflict of interest in research. In R.J. Porter and T.E. Malone (Eds.), *Biomedical research: Collaboration and conflict of interest* (pp. 121–162). Baltimore: The Johns Hopkins University Press.

Prien, R.F., and Levine, J. (1984). Research and methodological issues for evaluating the therapeutic effectiveness of antidepressant drugs. *Psychopharmacology Bulletin, 10,* 250–257.

Raskis, H.A., and Smarr, I.R. (1957). Psychopharmacotherapeutic research. *A.M.A. Archives of Neurology and Psychiatry, 2,* 202–209.

Robins, E., and Guze, S.B. (1970). Establishment of diagnostic validity in psychiatric illness: Its application to schizophrenia. *American Journal of Psychiatry, 126,* 107–111.

Rodin, G., and Voshart, K. (1986). Depression in the medically ill: An overview. *American Journal of Psychiatry, 143,* 696–705.

Ross, C.A. (1986). Biological tests for mental illness: Their use and misuse. *Biological Psychiatry, 21,* 431–435.

Ross, C.A. (1995). Errors in logic in biological psychiatry. In C.A. Ross and A. Pam (Eds.), *Pseudoscience in biological psychiatry* (pp. 85–128). New York: John Wiley and Sons, Inc.

Rutter, M. (1983). Behavioral studies: Questions and findings on the concept of a distinctive syndrome. In M. Rutter (Ed.), *Developmental neuropsychiatry* (pp. 259–279). New York: Guilford Press.

Sachs, O. (1987). *The man who mistook his wife for a hat.* New York: Harper and Row.

Sarbin, T.R. (1990). Toward the obsolescence of the schizophrenia hypothesis. *The Journal of Mind and Behavior, 11,* 259–284.

Schou, M. (1968). Lithium in psychiatric therapy and prophylaxis. *Journal of Psychiatric Research, 6,* 67–95.

Shea, M.T., Elkin, I., Imber, S.D., Sotsky, S.M., Watkins, J.T., Collins, J.F., Pilkonis, P.A., Beckham, E., Glass, D.R., Dolan, R.T., and Perloff, M.B. (1992). Course of depressive symptoms over follow-up: Findings from the NIMH treatment of depression collaborative research program. *Archives of General Psychiatry, 49,* 782–787.

Shear, K.M., Cooper, A.M., Klerman, G.L., Busch, F.N., and Shapiro, T. (1993). A psychodynamic model of panic disorder. *American Journal of Psychiatry, 150,* 859–866.

Shear, K.M., and Fyer, M.F. (1988). Biological and psychopathological findings in panic disorder. In A.J. Frances and R.E. Hales (Eds.), *Review of psychiatry, Volume 7* (pp. 29–53). Washington, D.C.: American Psychiatric Press.

Smith, A., Traganza, E., and Harrison, G. (1969). Studies on the effectiveness of antidepressant drugs. *Psychopharmacology Bulletin, 5,* 1–53.

Strauss, J.S. (1986). Psychiatric disorders: A reconsideration based on longitudinal processes. In T. Millon and G.L. Klerman (Eds.), *Contemporary directions in psychopathology: Towards the DSM-IV* (pp. 257–263). New York: The Guilford Press.

Strauss, J.S. (1992). The person — key to understanding mental illness: Towards a new dynamic psychiatry, III. *British Journal of Psychiatry, 161* (Supplement 18), 19–26.

Strauss, J.S., Gabriel, R.K., Kokes, R.F., Ritzler, B.A., VanOrd, A., and Tarana, E. (1979). Do psychiatric patients fit their diagnoses? *The Journal of Nervous and Mental Disease, 167,* 105–113.

Strauss, J.S., Hafez, H., Leiberman, P., and Harding, C.M. (1985). The course of psychiatric disorder III: Longitudinal principles. *American Journal of Psychiatry, 142,* 289–296.

Strub, R.L., and Black, F.W. (1985). *The mental status examination in neurology.* Philadelphia: F.A. Davis Company.

Stuss, D.T., and Benson, D.F. (1986). *The frontal lobes.* New York: Raven Press.

Summerfield, A. (1978). Behavioral toxicity — The psychology of pollution. *Journal of Biosocial Science, 10,* 335–345.

Taylor, R.L. (1982). *Mind or body.* New York: McGraw–Hill.

Teicher, M.H., Glod, C., and Cole, J.O. (1990). Emergence of suicidal preoccupation during fluoxetine treatment. *American Journal of Psychiatry, 147,* 207–210.

Teicher, M.H., Glod, C., and Cole, J.O. (1993). Antidepressant drugs and the emergence of suicidal tendencies. *Drug Safety, 8,* 186–212.

Terr, L.C. (1991). Childhood traumas: An outline and overview. *American Journal of Psychiatry, 148,* 10–20.

Thomson, R. (1982). Side effects and placebo amplification. *British Journal of Psychiatry, 140,* 64–68.

Vaillant, G.E. (1984). The disadvantages of DSM-III outweigh its advantages. *American Journal of Psychiatry, 141,* 542–545.

Vaillant, G.E. (1988). What is a case? *Archives of General Psychiatry, 45,* 313–319.

Vaughn, C.E., and Leff, J.P. (1976). The influence of family and social factors on the course of psychiatric illness. *British Journal of Psychiatry, 129,* 126–137.

Vega, W.A., and Murphy, J.W. (1990). *Culture and the restructuring of community mental health.* New York: Greenwood Press.

Wallerstein, R.S. (1986). *Forty-two lives in treatment.* New York: The Guilford Press.

Warnes, H. (1982). Physical illness in the psychiatric patient. In E.K. Korany (Ed.), *Physical illness in the psychiatric patient* (pp. 119–142). Springfield, Illinois: Charles C. Thomas, Publisher.

Weissman, M.M., and Klerman, G.L. (1980) Debate on psychiatric epidemiology. In reply [to L.Srole and A.K. Fisher]. *Archives of General Psychiatry, 37, 1423–1424.*

Wetzler, S., and van Pragg, H.M. (1989). Assessment of depression. In S. Wetzler (Ed.), *Measuring mental illness* (pp. 71–89). Washington, D.C.: American Psychiatric Press, Inc.

Widiger, T.A., and Shea, T. (1991). Differentiation of Axis I and Axis II disorders. *Journal of Abnormal Psychology, 100,* 399–406.

Wilson, M. (1993). DSM-III and the transformation of American psychiatry: A history. *American Journal of Psychiatry, 150,* 399–410.

Wing, J.K. (1987). The long-term social adaptation in schizophrenia. In N.E. Miller and G.D. Cohen (Eds.), *Schizophrenia and aging* (pp. 183–188). New York: The Guilford Press.

Zinberg, N.E. (1976). Observations on the phenomenology of consciousness change. *Journal of Psychedelic Drugs, 8,* 59–76.