
MARK B. SOBELL† and LINDA C. SOBELL
Clinical Institute, Addiction Research Foundation, 33 Russell Street, Toronto, Ontario M5S 2S1 and Department of Psychology, University of Toronto, Toronto, Ontario M5S 1A1, Canada

(Received 31 October 1983)

Summary—A point-by-point response is presented to Pendery et al.’s (1982) critique of the study “Individualized Behavior Therapy for Alcoholics (IBTA)”. Two independent, external investigations have critically examined the original records from the IBTA study. The IBTA study compared two different techniques for treating poor-prognosis, chronic alcoholic state hospital patients. It was found that a broad-spectrum behavioral treatment oriented to controlled drinking was a more effective treatment than was the standard hospital treatment program based on conventional wisdom. The Pendery et al. critique gave the appearance of being a refutation by presenting findings for only one group of Ss in a comparative study. In this response, it is shown that the experimental and control Ss were justifiably classified as gamma alcoholics, that Ss were randomly assigned to groups, and that the two groups were comparable in terms of pretreatment characteristics. Moreover, as regards the originally reported 2-yr treatment outcome findings, it is shown that Pendery et al. reported no specific events which were not already documented in the original study records which formed the basis for publications about the IBTA study. Finally, even in terms of long-term outcomes, i.e. mortality rates 10-11 yr after treatment, the experimentally-treated group (20% mortality) continued to fare better than the traditionally-treated group (30% mortality). Ironically, rather than Pendery et al.’s report being a refutation of the original published accounts of the IBTA study, it is concluded that the data reported by Pendery et al. actually strengthen the validity of the original publications. The attack on the IBTA study can be meaningfully viewed as a reflection of the scientific revolution presently underway in the alcohol field. The central impetus for conceptual change in the alcohol field, however, derives less from controlled-drinking research than from the lack of empirical support for conventional wisdom. Ideas are changing because the traditional view no longer inspires progress.

INTRODUCTION

An article authored by Pendery, Maltzma and West (1982) and published in the 9 July 1982 issue of Science magazine has purported to reexamine the findings of our study of Individualized Behavior Therapy for Alcoholics (IBTA; Sobell and Sobell, 1972, 1973a, b, 1976, 1978), for which follow-up results were reported in Behaviour Research and Therapy (Sobell and Sobell, 1973b, 1976). We were not offered an opportunity by Science to simultaneously publish a response, or even informed that a manuscript highly critical of our research had been under review for some time. Moreover, publication of the Pendery et al. (1982) article was accompanied by massive news media coverage generally portraying the findings of the IBTA study as invalid. In this article, we respond to the issues raised by the Pendery et al. (1982) article on a point-by-point basis.† We conclude that rather than the Pendery et al. (1982) article being a refutation of our findings, their data actually strengthen the validity of our original reports and conclusions. Moreover, the attack on our study, while unprecedented in its intensity, is quite consistent with Kuhn’s (1970) philosophy of scientific revolutions, a process that is exemplified by recent events in the field of alcohol studies.

Some readers may find it puzzling that a response to a scientific critique should appear in a different journal than that in which the critique was published. The reason for this is quite simple. After the publication of the Pendery et al. (1982) article, we were informed by Science that we could

*The Addiction Research Foundation requires that all publications carry the following statement: the views expressed in this publication are those of the authors and do not necessarily reflect those of the Addiction Research Foundation.
†Requests for reprints may be addressed to either author at the Clinical Institute.
‡In their article, Pendery et al. (1982) also took issue with a third year outcome study of the IBTA Ss independently conducted by Caddy, Addington and Perkins (1978). Since we were not involved in the actual conduct of the third year independent follow-up study, except for informing Ss about it, it would not be appropriate for us to respond to issues regarding that study.

413
submit for review for publication, a technical comment on the issues raised in that article, but that our response would have to be limited to 2000 words. In the context of all the events surrounding this controversy, it was our opinion that a more thorough examination of the issues was warranted and necessary. Thus, publication in a journal in which our original research had appeared was a preferred alternative.

Since the Pendery et al. (1982) critique was published in a different journal than our response, it is incumbent upon us to summarize the major issues. At the outset, however, we wish to emphasize that readers who wish to study this controversy in greater detail should consult the original publications, including those which we have authored. In fact, as we will show, if this is done conscientiously the reader might wonder exactly what is in dispute. This raises the issue of the greater context of the present conflict, which deserves some mention before proceeding to a discussion of the controversy over the IBTA study.

**THE CONTROVERSY IN CONTEXT**

The notion that some individuals with alcohol problems may recover fully or improve significantly without abstinence was not novel in the early 1970s when the IBTA study was conducted. In fact, a major foundation of the IBTA study was that there had been repeated reports in the scientific literature of persons who had recovered and were able to drink without incurring serious problems (Sobell and Sobell, 1972, 1973a). Such reports have multiplied considerably in recent years, perhaps due to investigators designing studies so as to allow measurement of the phenomenon. This massive body of evidence has been comprehensively and critically reviewed by Heather and Robertson (1981, 1983). Reports such as these contradict conventional wisdom in the alcohol field, as noted by Pendery et al. (1982). However, this would not be the first case where conventional wisdom was found to be more convention than correct. It is also important to recognize that conventional wisdom in the alcohol field, especially with respect to treatment, lacks a strong empirical basis (Pattison, Sobell and Sobell, 1977; Heather and Robertson, 1981), even though it is firmly entrenched among rank-and-file workers and among many scientists. Consequently, the reaction to reports of nonabstinent recoveries has at times been virulent.

One of the earliest and best known examples of such a reaction was the unprecedented barrage of criticism (see Davies, 1963), much of it irrational, that surrounded the late D. L. Davies’ (1962) published report of “Normal drinking in recovered alcohol addicts”. The other most widely publicized attack occurred with the publication of the first (Armor, Polich and Stambul, 1978) and second (Polich, Armor and Braiker, 1981) Rand Reports. Heather and Robertson (1981) present a thorough discussion of these debates, and Armor et al. (1978, Appendix B) present extensive documentation of the conflict over the first Rand Report. Also, the July 1980 issue of the *Journal of Studies on Alcohol* contains extensive commentary on the findings of the second Rand Report (Beauchamp et al., 1980). It is of more than passing interest, however, that the senior author of the Pendery et al. (1982) article played a significant role in disputing the Rand findings (see Armor et al., 1978, Appendix B). This, then, is the background of the present controlled-drinking controversy, which is but the most publicly visible part of the greater paradigm conflict that embroils the alcohol field (Pattison et al., 1977; Heather and Robertson, 1981; Marlatt, 1983; Moos and Finney, 1980; Beauchamp, 1980; Polich et al., 1981).

A second major precipitant of the present scientific crisis in the alcohol field has been the failure of traditional ideas to provide effective treatments for alcohol problems. However, this aspect of the crisis has received little attention, particularly from advocates of the conventional wisdom. Beneath the rhetoric, though, one finds that traditional approaches, including even Alcoholics Anonymous, have little empirically demonstrated effectiveness, particularly for the type of S population in the IBTA study (Miller and Hester, 1980; Emrick, 1974, 1975, 1982; Costello, Biever and Baillargeon, 1977; Gordis, Dorph, Sepe and Smith, 1981; Backeland, 1977; Edwards, 1980; Polich et al., 1980; Vaillant, 1983). This fact is well known and has occasionally been commented upon by more objective advocates of conventional wisdom (e.g. Vaillant, 1983). Therefore, since it is directly relevant to the IBTA study, it is important to bear in mind that the controlled-drinking controversy has become the central focus of a paradigm conflict involving a broad challenge to traditional conceptualizations of alcohol problems. An often unspoken but important underlying
reason for much of the controlled-drinking research is the marked absence of effective treatments for this widespread public health problem.

Thus, even today it is the case that there is no known panacea for alcohol problems, and particularly no scientifically-validated effective treatment for chronic alcoholics who enter public treatment programs. In conducting treatment research on the latter population, the emphasis is on comparing the relative effectiveness of different treatment approaches. In this sense, an analogy can be made to cancer research. Treatment outcomes are a mixed bag, heavily speckled with morbidity and even mortality. Nevertheless, some treatments are associated with better outcomes than others. Such a comparative evaluation was the core research question of the IBTA study. The examination of long-term outcomes for only one group in a comparative study of treatments for chronic alcoholics (any comparative study) can easily be portrayed as tragic, as exemplified in the approach taken by Pendery et al. (1982). However, for scientific purposes such an evaluation is meaningless. It fails to address the fundamental research question: is one type of treatment more effective than another? And in failing to address this question, it shields from discovery a major reason why the paradigm conflict exists—traditional treatments for chronic alcoholics in publicly-funded treatment programs have little demonstrated efficacy. That is a central reason why the investigation of alternative treatment approaches was, and remains, important; and the finding that nonconventional methods of treatment may yield superior outcomes gives further impetus to the process of paradigm change. As Kulik (1970) postulated, the process of paradigm change is fraught with conflict that goes beyond debates of scientific merit.

BACKGROUND

The IBTA study involved four groups of Ss (N = 70) who were inpatients in the alcoholism treatment program at Patton State Hospital (Calif.) in 1970 and early 1971. The Pendery et al. (1982) critique was almost exclusively concerned with the outcome for one group, the group designated as Controlled-drinker Experimental (CD-E) in our publications and as Controlled-drinking Ss by Pendery et al. (1982). However, the IBTA study was designed to include two, related two-group comparisons. Patients who volunteered for the study were first screened and assigned by the research staff to one of two goal-eligibility conditions: either a controlled-drinking goal or a nondrinking (abstinence) goal. Criteria used in making goal-eligibility decisions were conservative and have been discussed elsewhere (Sobell and Sobell, 1972, 1973a). All Ss were evaluated as having a history of physical dependence on alcohol (gamma alcoholics in Jellinek's, 1960, terminology). Within each goal-eligibility group, Ss were then randomly assigned to either the traditional, abstinence-oriented treatment program offered by the hospital, or to an experimental treatment program (in combination with the regular hospital treatment program). Thus, patients selected for the nondrinking goal-eligibility condition were randomly assigned to either an experimental (Nondrinker Experimental, ND-E, n = 15) or control (Nondrinker Control, ND-C, n = 15) group, both involving a treatment goal of abstinence but differing in treatment procedures. Likewise, patients selected for the controlled-drinking goal-eligibility condition were randomly assigned to either the regular, abstinence-oriented hospital treatment program (Controlled-drinker Control, CD-C, n = 20), or to an experimental, controlled drinking-oriented treatment program in conjunction with all components of the regular hospital treatment program except the abstinence goal (Controlled-drinker Experimental, CD-E, n = 20). To avoid confusion, and because this article deals almost exclusively with the controlled-drinking goal-eligible Ss, we will henceforth refer to group CD-E as the 'experimentally-treated' group, and to group CD-C as the 'traditionally-treated' group.

Both the 17-session experimental treatment and the regular hospital treatment program have been described in our earlier publications (Sobell and Sobell, 1972, 1973a) and will not be described here, as the details are not relevant to the Pendery et al. (1982) critique. However, it should be noted that the experimental treatment differed from the traditional treatment in two major ways:

1) the experimentally-treated Ss participated in an intensive broad-spectrum behavioral treatment (as did the ND-E Ss); and
2) their treatment had an objective of controlled drinking.
Thus, as clearly stated in our original publications, any difference between these two groups could have related to the behavioral treatment, to the controlled-drinking goal, or both. Moreover, the differential contributions of these factors to outcomes cannot be determined from the design used.

Follow-up data were presented in extensive detail for the 2-yr period following hospital discharge, and complete data were reported for 69 of the 70 Ss in the study, including all 20 experimentally-treated Ss (CD-E) and 19 of the 20 traditionally-treated Ss (CD-C). Our major conclusions of present relevance were:

1. that the experimentally-treated Ss fared significantly better than the traditionally-treated Ss over the 2-yr follow-up interval;
2. that the experimentally-treated Ss, as confirmed by collateral informants, engaged in more days of abstinence and fewer days of heavy drinking than the traditionally-treated Ss; and
3. that only the experimentally-treated Ss engaged in a substantial amount of limited, nonproblem drinking during the follow-up.

For the second year of follow-up, individual S profiles were presented listing the number of days that each S spent in various drinking disposition categories (Sobell and Sobell, 1976), and for both years of follow-up individual S as well as group results were reported (Sobell and Sobell, 1973b, 1976). It is important to note that we found that Ss in all four groups in the study showed improvement over time, and that many Ss in all of the groups had a substantial number of dysfunctional days early in the course of follow-up. Such a pattern of recurrence of drinking problems early in follow-up, followed by improvement, has been found for other effective social-learning-based treatments for chronic alcoholics (Marlatt, 1983).

The Pendery et al. (1982) Critique

The Pendery et al. (1982) critique of the IBTA study was multifaceted. The major issues will be summarized here and then elaborated in greater detail along with our response. Since issues such as these are at times open to various interpretations, readers who wish a first-hand acquaintance with the conflict should consult the original publications.

The Pendery et al. (1982) critique, as relevant to our response, had three major areas of focus:
1. it raised questions about the design of the study as enacted (specifically, the nature of the S population and the comparability of the groups);
2. by inference it appeared to present follow-up findings for the initial 2 yr of follow-up for the experimentally-treated Ss which differed from those we had reported; and
3. it presented information on long-term treatment outcomes over an approx. 11-yr interval for only the experimentally-treated Ss. Pendery et al.'s (1982) article was reported to be based on interviews with Ss and collateral information sources (with initial contact "established in the period 1976 to 1979", p. 172), S affidavits and documentary data (e.g., hospital records). They stated that on this basis they had reached "conclusions that are very different from the conclusions of the Sobells" (p. 172).

Since self-reports of Ss and collaterals formed part of the Pendery et al. (1982) data set, as well as part of our own data set, the possibility that demand characteristics may have influenced these responses cannot be discounted, as we first discussed several years ago (Sobell and Sobell, 1973b). In this regard, the purpose of the Pendery et al. (1982) study assumes some importance, since they do not maintain that their interviews were conducted under conditions of interviewer blindness to their study's hypothesis. Although we did not become informed of it until December 1976, the hypothesis of the Pendery et al. study, as stated in the project protocol approved on the 5 May 1973, was:

"The hypothesis of the proposed project is that individualized behavior therapy for controlled drinking has not produced the effects reported by the investigators." (Pendery, Maltzman and Digan, 1973)

This, it must be noted, was well before Pendery et al. had established contact with any of the Ss.

Results of Two Independent Investigations

When we became aware of the pending publication of the Pendery et al. critique, we requested that our present employer commission an independent, external inquiry concerning the IBTA

MARK B. SOBELL and LINDA C. SOBELL
study. Such an inquiry was established in June 1982. The Committee was chaired by Dr Bernard M. Dickens (Ph.D—criminology, LL.D—medical jurisprudence), Professor of Law, Criminology and Community Health at the University of Toronto. The other Committee members were Dr Anthony N. Doob (Ph.D—psychology), Professor of Psychology and Director of the Centre of Criminology at the University of Toronto, Dr O. Harold Warwick (M.D.), Professor Emeritus and past Vice-president of Health Sciences at the University of Western Ontario, and Dr William C. Winegard (Ph.D—metallurgy), past President of the University of Guelph (Ontario). The Committee submitted its 123-page report* in late October 1982, in which is concluded that:

"The Committee finds there to be no reasonable cause to doubt the scientific or personal integrity of either Dr Mark Sobell or Dr Linda Sobell." (Dickens, Doob, Warwick and Winegard, 1982, p. 109)

For ease of reference, the Committee will henceforth be referred to as the Dickens Committee.

The Dickens Committee's Report has only partial relevance to the present article. In the main, it represents an independent confirmation of the existence of the research records from the IBTA study (see pp. 35–40 of the Dickens Committee's Report for a topical listing). The major records include original data sheets, complete follow-up records, arrest records, drivers records, audiotape recorded interviews with Ss, correspondence to and from Ss and telephone logs of follow-up contacts. In addition, the Committee received and verified written statements from all 8 individuals who were formerly employed as staff on the IBTA study. Other records pertinent to the in-hospital conduct of the study were left at Patton State Hospital in June 1971, when we moved to other positions, and were subsequently destroyed by hospital staff without our knowledge.

Our records were subsequently independently confirmed and cross-checked in the course of a Congressional inquiry conducted by the Subcommittee on Investigations and Oversight, Committee on Science and Technology of the United States House of Representatives. In a letter informing us of the results of the inquiry, the investigator wrote:

"My review of all available supports the findings of the Commission convened by the Addiction Research Foundation (also known as the 'Dickens' Commission') and fully supports their conclusion." (James E. Jensen, personal communication, 23 March 1983)

It should be noted that we cooperated fully with both of these inquiries and that the Congressional investigator had access to Congressional subpoena power. As the Dickens Committee aptly stated, however:

"No amount of data, notes or attestations would convince an atheist that God exists, or a believer that He does not" (Dickens et al., 1982, p. 40)

While this is undoubtedly true, we also believe that those who are not fully ideologically committed in the present conflict will agree that there has been independent confirmation of our records.

**RESPONSE TO THE CRITIQUE**

**Experimental Procedures**

**Subject characteristics**

A major contention of Pendery et al. (1982) was that 4 of the 20 experimentally-treated Ss did not, in Pendery et al.'s view, qualify as gamma (physically dependent) alcoholics at the time of their participation in the IBTA study. Pendery et al.'s (1982) conclusion was based on S's self-reports elicited several years after their participation in the IBTA study, and presumably (at least in one case) on Patton State Hospital records. For 1 of the 4 Ss (CD-E 18) more extensive documentation was presented, consisting of confirmation of the S's self-report by his wife, Patton records and his score of zero "to all relevant questions" (Pendery et al., 1982, p. 175) on the Alcohol Dependence

*Copies of the Dickens Committee's Report can be obtained at a per copy cost of $7.00 (Can.) for American orders and $4.00 (Can.) for Canadian orders from: Information Centre, Addiction Research Foundation, 33 Russell Street, Toronto, Ontario M5S 2S1, Canada (limited supply).
Scale (ADS; Stockwell, Hodgson, Edwards, Taylor and Rankin, 1979). That this individual should require special consideration was presumably important because Pendery et al. (1982) were forced to conclude that this S had "successfully maintained his pattern of controlled drinking" (p. 174) throughout the long-term follow-up period.

In their commentary on this issue, the Dickens Committee chose to point out the difficulty involved in making reliable diagnoses of gamma alcoholism, especially since our original data indicated that the 4 Ss in question had all experienced only psychomotor tremors as withdrawal symptoms. Indeed, Caetano, Edwards, Oppenheim and Taylor (1978), in a study of interrater reliability, found that "the items which proved most unreliable was the Jellinek classification" (p. 193), with 5 trained raters (4 psychiatrists, 1 psychologist) agreeing in only 30% of cases. This was in comparison to interrater agreement rates above 70%, on most other items. Similarly, in a recent evaluation of the 2-week test-retest reliability of the Severity of Alcohol Dependence Questionnaire (a very slightly modified variant of the ADS), Stockwell, Murphy and Hodgson (1983) found that item reliabilities ranged from \( r = 0.35 \) to \( r = 0.82 \), and the total scale reliability was \( r = 0.85 \). These findings indicate that the scale has good reliability, but that response variability is found even with as short a test-retest interval as 2 weeks. In view of the fact that S background and drinking-history variables did not differ significantly between the two groups (Sobell and Sobell, 1972, 1973a), and if anything, showed the experimentally-treated Ss to have had a slightly more severe history than the traditionally-treated Ss, the Dickens Committee concluded:

"The concurrence rate of 80 percent in the Sobells' and the Pendery et al. assessment of the 20 CD-E subjects may be as high as can reasonably be expected, given judgments of less than perfect reliability." (Dickens et al., 1982, p. 74)

Although the Dickens Committee's approach to this issue may be a sufficient resolution, other more direct evidence also supports our classification of the 4 Ss as physically dependent. Before discussing that evidence, we should point out, with regard to Pendery et al.'s (1982) data, that little is known about how the passage of time (in this case, several years), as well as interviewer bias, might influence retrospective self-reports of symptoms such as psychomotor tremors. As well, it is important to recognize that our designation of Ss as physically dependent in 1970-1971 was based on three data sources. First, the S must have reported withdrawal symptoms to the research staff in a screening questionnaire and/or interview for the study. Second, as part of the medical screening for eligibility to participate in the study, the alcoholism treatment program director, a psychiatrist, must have concluded that the S could be appropriately classified as a gamma alcoholic (the psychiatrist who made these decisions is now deceased). Third, the Patton medical records were reviewed for evidence of withdrawal symptoms or history. However, we viewed the Patton records as extremely unreliable and merely as one possible source of data. This was based on our familiarity with the records at the time, but our position has since been reinforced by the two independent inquiries. We have in our possession, a few copies of documents from the Patton records which we were permitted to obtain years ago to assist us in locating Ss for follow-up, and in one case (a CD-C S) we possess extensive documentation. A review of these documents revealed glaring errors. For example, in the one case where the documentation includes a medical and psychiatric history, all information was obtained during a single interview when the S was intoxicated on paraldehyde and 'suffering from alcohol withdrawal'. Errors ranged from relatively benign points, such as an incorrect place of birth and age, to incorrect histories of past hospitalizations and arrests. In the hospital record the S was reported as having had no past hospitalizations, and two minor arrests. However, a record check which we later conducted revealed one past alcohol-related hospitalization (which the S had accurately reported to us in the research screening interview) and three serious unreported arrests. In other cases, for example, Ss are reported as participating in research when they did not, as not participating in research when they did, and blank, yet signed and physician-witnessed, consent for treatment forms were included in the patient's hospital record. In all, it would appear that we, as well as subsequent researchers at Patton State Hospital (e.g. Baker, Udin and Vogler, 1975; Vogler, Compton and Weissbach, 1975) were well justified in not relying on the hospital records as a primary data source.

As concerns the other evidence supporting our designation of all experimentally-treated Ss as gamma alcoholics, the most important consists of tape-recorded statements made by the Ss
themselves. As part of the follow-up of the IBTA study, 67 of the 68 living Ss were administered a lengthy standardized, audiotape-recorded interview at the completion of follow-up. The interview format was designed in 1971 and was oriented toward illuminating those aspects of the treatment program, follow-up procedures and change following treatment that could not be adequately reflected in the quantitative measures of outcome (Sobell and Sobell, 1976, 1981). Because the vast majority of interviews were completed before Pendergy et al. had expressed any interest in the IBTA study, the interview naturally did not directly address questions subsequently raised by Pendergy et al. (1982). It was not designed for that purpose. On the issue of diagnostic classification of Ss, for instance, the interview did not inquire about past withdrawal symptoms. This was because that information was already on file. However, as a way of attempting to characterize the extent of change from pretreatment to posttreatment, Ss were asked to describe their drinking behavior prior to entering the research program. Their responses, as well as arrest data, are directly pertinent to the issue of gamma classification, because they are highly consistent with that classification, and because they were obtained before any questions were raised about the study. These clinical data were obtained from the 4 Ss in question during the period from October 1972 through March 1973 (within 7 days of the end of each S's second year of follow-up). Persons with clinical experience in the alcohol field will recognize that the interview responses are consistent with a diagnosis of gamma alcoholism.

Subject CD-E 18. This S's arrest record indicates two alcohol-related arrests in 1966, one for drunk driving and one for public drunkenness. Relevant excerpts from the second year interview are as follows.

Interviewer (I): Why did you initially come to Patton State Hospital 2 years ago?
Subject (S): Well, I was past the point of helping myself in controlling my drinking, so I had to have help from somebody a little smarter than I was.
I: Now the next series of questions are about any changes in your drinking habits over the last few years. Would you say that your drinking has increased or decreased?
S: Well, I'm drinking very considerably lot less than I was before.
I: OK, could you describe how much you were drinking before you went into Patton on the average?
S: Prior to going into Patton the last year I was drinking probably daily or at least 4 or 5 days a week heavy.
I: And how much is heavy?
S: Well, probably from a pint to a quart of whiskey a day, besides the beer.
I: Has the time of day that you do your drinking changed?
S: Yes, I never drink in the mornings anymore, never in the a.m.
I: OK, would you describe yourself before you went into the hospital as a binge, periodic or continuous drinker?
S: Well, I was both (sic). There would be weeks, you know week after week, and I would drink every day and heavy, and then I would go for a week or 2 weeks and not drink anything. Then a problem would come, and I think well I'm going to get drunk and that's what I'd do.
I: Has the environment where your drinking occurs changed? Where did you use to drink two years ago?
S: I used to drink at (name deleted). You know I'd go out there and get plastered, and then stop at (name deleted) and have a lot of drinks.
I: Do you believe you have or have had an alcoholic drinking problem?
S: I know I had an alcoholic drinking problem, there was no doubt about it.

Those familiar with the Alcohol Dependence Scale (Stockwell et al., 1979), administered retrospectively to S CD-E 18 by Pendergy et al. several years after the IBTA study, will recognize that this S's 1972 interview responses are not at all consistent with scoring a zero even on a nonmodified version of that scale, since the scale includes several items on amount of alcohol consumption and also on morning drinking. However, it should be noted that Pendergy et al. (1982) only reported that the S scored a zero "on all relevant (italics added) questions" (p. 175). It should be further noted that the amount of daily consumption reported by this S is consistent with that
required to produce physical dependence in studies of experimental intoxication (e.g. Gross, Lewis and Hastey, 1974) as well as that reported by individuals clinically assessed as physically dependent on alcohol (e.g. Shaw, Kolesar, Sellers, Kaplan and Sandor, 1981). These clinical data, therefore, are consistent with our records of the S having reported experiencing psychomotor tremors as a withdrawal symptom when assessed in 1970. Finally, in our tape-recorded interview, S CD-E 18 reported himself as having lost control over his drinking, a cardinal feature of gamma alcoholism.

Subject CD-E 17. This S's arrest record indicates one arrest for drunk driving in 1969. Relevant interview excerpts are as follows.

I: Has your drinking increased or decreased over the last 2 years since before you went into the hospital?
S: Decreased.
I: Could you tell me how much you were drinking before you went into the hospital?
S: Well, I can't tell how much. It was a lot.
I: Can you give me a guess of what is a lot?
S: Every other day, all day.
I: OK, what kind of liquor were you drinking?
S: Usually beer.
I: If you had to give a number of cans that you were drinking before you went into the hospital, how many would you say you would have done on an average day?
S: 20.
I: Would you describe yourself 2 years ago as a binge, periodic or continuous drinker?
S: Continuous.
I: Has the environment where your drinking occurred changed? That is, where did you use to do most of your drinking 2 years ago?
S: Well, yes, that has changed. It's at home.
I: It's at home now?
S: Yes.
I: Where did you use to do it?
S: In bars.
I: Do you believe you have or have had an alcoholic drinking problem?
S: Well, I believe I had one.


I: Would you say that your drinking has increased or decreased in the last 2 years compared to before you went into the hospital?
S: Only one way, and that's been down.
I: OK, would you tell us how much you drank before you went into the hospital on an average day?
S: Oh, I would say on an average a day, well just, you know, prior to the time I came in?
I: Right.
S: I would say on an average a day I'd probably, in between sneaking off to bars, this, that and the other thing, it wouldn't surprise me if I didn't put away a good fifth of whiskey a day.
I: What kind of liquor did you drink before?
S: Oh, I was always getting into trouble on whiskey.
I: Has the time of day changed that you do your drinking?
S: No.
I: OK.
S: Oh, do you mean from before?
I: From before, yes.
S: Oh, good Lord yes. Before I would even have little half pints of whiskey hid. And you know you get up in the morning and drink.
I: Would you describe yourself 2 years ago as a binge, periodic or continuous drinker?
S: Well, I would consider it binge–periodic, but too heavy of drinking in between.
I: Where did you use to do your drinking and with whom?
S: I used to go off by my lone and sit in a corner of a bar or in a dark corner in a cocktail lounge and get stoned.
I: Do you believe you have or have had an alcoholic drinking problem?
S: Oh, I know I had a drinking problem. Let's face it, or I wouldn't have been in jail so many times.

Subject CD-E 20. This S's arrest record indicates four previous alcohol-related arrests, all for drunk driving (two in 1966, one in 1969, one in 1970). Relevant interview excerpts are as follows.

I: Would you say that your drinking has totally increased or decreased?
S: It has decreased.
I: OK, could you tell us how much you used to drink 2 years ago on the average?
S: Oh, I can't remember, you know maybe a pint, a fifth, as much as I could get.
I: What kind of liquor?
S: Primarily scotch.
I: Has the time of day that you do your drinking changed?
S: Yeah.
I: What time did you used to do it before?
S: All day, anytime of day.
I: OK, would you say that your drinking 2 years ago was binge, periodic or continuous?
S: It was continuous.
I: Do you believe you have or have had an alcoholic drinking problem?
S: Yes, I feel that I have had a problem.

Summary. To reiterate what is at issue, Pendery et al. (1982) asserted that in their opinion 4 of the 20 experimentally-treated Ss should not have been classified as gamma alcoholics. Their main sources of information were hospital records of questionable accuracy and Ss' retrospective self-reports gathered several years after their participation in the IBTA study. Our classification of Ss was based on their self-reports at the time of the study, an assessment by a psychiatrist and information available in hospital records. As noted by the Dickens Committee, diagnoses of gamma alcoholism cannot be made with high reliability, and the two groups of Ss were comparable in terms of background and drinking history data. In fact, if any difference existed between the groups it was that the experimentally-treated Ss had more severe drinking problem histories. As concerns the 4 Ss claimed by Pendery et al. (1982) as having been inappropriately classified as gamma alcoholics, the interview data presented above, excerpted from interviews conducted in 1972 and early 1973, is consistent with our classification of the Ss as gamma alcoholics. Considering the totality of evidence, we conclude that the issue of diagnostic classification is moot.

Random assignment to groups

It has already been discussed that the experimentally-treated and traditionally-treated Ss did not differ significantly in terms of background and drinking history variables. Likewise, our basis for classifying all experimentally-treated Ss as gamma alcoholics has been discussed at length. The further question of group comparability raised by Pendery et al. (1982) relates to the temporal ordering of Ss within the study groups. Specifically, Table 1 of their article (p. 171) lists the Ss by date of admission to Patton State Hospital, and Pendery et al. (1982) noted that the majority of experimentally-treated Ss were admitted to the hospital prior to the traditionally-treated Ss. This raises the question of whether the Ss were, indeed, randomly assigned to treatment conditions.

The Ss were, without exception, randomly assigned to groups. When the IBTA study was initiated, Ss were assigned to treatment conditions on the basis of a simple coin flip. The use of
such a procedure was quite common at that time. Efron (1971), for example, noted that:

"In order to avoid biasing the results of the experiment it is customary to make the assignments by independent flips of a fair coin. but in small-sized experiments this may result in a severe imbalance between the number of treatments and controls."

(p. 403)

Such a problem of imbalance developed during the course of the IBTA study, and as the study continued, it became evident that the sample size would be too small to result in equal-sized groups unless the procedure was modified. Thus, balancing was accomplished by the use of a weighted random-assignment procedure. This entailed waiting until a group of Ss had volunteered and been medically cleared for the study, and then assigning them to treatments on a random but disproportionate basis. The weighted random assignment was also performed by coin flip (using multiple coins), and like the other coin flips was always witnessed. We were able to reconstruct for the Dickens Committee and the Congressional inquiry the actual procedure followed. Moreover, all of the former research staff have testified to the use of the randomization procedure. For example, the following are excerpts from some of the former staff members’ statements.

*Staff member A*

"After the subject had been assigned to one of the two (goal) groups then a flip of the coin was used to randomly assign them to the Control Group or Experimental Group. Once assigned a subject would always remain in that group throughout the life of the research program. To my knowledge this procedure of randomly assigning subjects to Control Groups and Experimental Groups was never varied in the 1 1/2 years that I participated in the Alcoholic Research Program."

*Staff member B*

"Without exception, assignment to experimental and control groups was carried out according to this random process. There was never an instance of group placement on any other basis, nor was there ever an instance of changing group assignment after the coin toss. These things I remember clearly, because it sometimes seemed unfair that the arbitrariness of assignment excluded patients who wanted to be experimental subjects."

*Staff member C*

"It is important for me to note that I was present during a great many of the patient interviews and coin tosses and I am certain that assignment to experimental and control conditions was done in a truly random fashion."

*Staff member D*

"In every instance the selection for the control group or the experimental group was done randomly and the data reported was the actual data collected from these two random groups."

*Staff member E*

"This (experimental or control group) determination was made by a coin toss. Once a patient had been placed in the control group or the program group, that designation was never switched. I cannot recall any instance where a patient was changed from one group to the other."

It should be noted, as well, that date of admission to the hospital does not necessarily bear a direct relationship to order of entry into the study. Data on actual dates of participation in the study were apparently among those which hospital personnel destroyed some time after we had left employment at the hospital. Since the assignment was in fact randomly determined, a detailed accounting of the assignment procedures was not included in our published reports. However, a careful reading of our 1973 original account of the IBTA study reveals that, 9 yr prior to the Pendery et al. (1982) article, we had reported that the experimentally-treated Ss tended to have been discharged from the hospital earlier than the traditionally-treated Ss (Sobell and Sobell, 1973a, Table 4, p. 63).

Given the randomization procedures used, the remaining question is whether those procedures might have inadvertently introduced a systematic difference between groups. Such a bias would
presumably have related to seasonal differences among Ss. Such differences would not have been expected, however, because Ss for these groups were prescreened by stringent criteria, such that they had greater social stability than the majority of alcoholic patients in the hospital and they were not transients. The lack of pretreatment differences between groups on demographic and drinking problem history variables underscores that the groups were comparable.

Summary. Ss were randomly assigned to the experimentally-treated and conventionally-treated groups in the IBTA study. At first, assignment was determined by a simple, equal probability coin flip. Later, in order to equate group size, a weighted randomization procedure was used. All former research staff submitted statements to the Dickens Committee affirming the randomization procedures. Moreover, the groups did not differ significantly in terms of background variables and drinking problem histories; if anything, the experimentally-treated Ss had more severe drinking problems prior to the IBTA study.

2-yr Posttreatment Outcomes

Pendery et al.'s (1982) reasons for not presenting comparative outcome data were threefold. First, they reasoned, the traditionally-treated Ss were aware that they had been eligible for the experimental treatment, and this would confound any interpretation of effects. Indeed, we (Sobell and Sobell, 1976, 1978) raised this question several years prior to Pendery et al. (1982). In some ways, the informed-consent procedures used in the IBTA study in 1970 and 1971 were ahead of their time, which probably related to the fact that California was one of the first states to attend to the protection of research Ss' rights. Such informed consent is now the model for research studies, and while it may at times make interpretation of results more difficult, we are not aware of any scientific evidence showing consent procedures to have a practical effect on outcomes in alcohol treatment research. As discussed earlier, it should be kept in mind that there is presently little evidence that even massive interventions have much effect on the drinking of chronic alcoholics, such as the Ss in the IBTA study. Furthermore, a major reason why Ss were assigned (if not contraindicated by other factors) to the controlled-drinking goal-eligibility condition was that they were not accepting of an abstinence goal. Finally, it is important to note that on the basis of the audiotape-recorded interviews conducted with Ss at the end of their second year of follow-up, 42% of the traditionally-treated controlled-drinking goal-eligible Ss (CD-C group) responded that they had at some time felt rejected or resentful about not being allowed to participate in the experimental treatment, whereas 43% of the traditionally-treated abstinence goal-eligible Ss (ND-C group) expressed similar feelings (Sobell and Sobell, 1978). If the treatment goal condition had exerted a strong effect, as one might surmise from the Pendery et al. (1982) article, then it would be expected that proportionally more of the CD-C than ND-C Ss would have reported resentment. In any event, this criticism, already noted by us in the literature, would not justify the presentation of outcome data for only one group in a comparative study.

The second argument for avoiding the presentation of comparative data given by Pendery et al. (1982) was that the groups may not have been comparable. This contention, as has been discussed at length, is erroneous.

The final reason stated by Pendery et al. (1982) for not presenting comparative outcomes was that they were not addressing such a comparison, but rather "whether controlled drinking is itself a desirable treatment goal" (p. 172). Before proceeding, it is of major importance to note that Pendery et al. (1982) apparently conceded a poor outcome for the traditionally-treated group. To use Pendery et al.'s (1982) own words, they characterized the outcomes of the traditionally-treated group as ones that "all agree fared badly" (p. 173). As should be clear by now, the outcome of any single group in a comparative study of treatment effectiveness with chronic alcoholics of poor prognosis is meaningless. Even worse, it can be used to discuss results out of context and to justify unwarranted conclusions. In fact, such an approach would portray the outcomes for the traditionally-treated Ss as catastrophic. As will become clear shortly, and is already clear to anyone who has read our publications with care, we never presented our findings in terms of Horatio Alger-type outcomes. Rather, we contended that one group of Ss fared significantly better than another over a 2-yr period following treatment. This was a direct inference based on our data, and therefore the basis for and validity of our findings deserves discussion. This lead to the second
major focus of the Pendery et al. (1982) critique; namely, their conclusion that their 2-yr outcome results differed from those which we reported.

Hospitalizations

One of the most perceptually striking findings presented by Pendery et al. (1982) concerned hospitalizations incurred by Ss during the first 2 yr of follow-up. A table, which occupied more than 10% of the space for the entire article, presented rehospitalization data for only the experimentally-treated Ss "within approximately 1 year of their participation in the research project" (Pendery et al., 1982, Table 2, p. 172). Pendery et al. (1982) neither claimed nor disclaimed that their data were inconsistent with our own, but only contended that "the references to hospital and jail incarcerations in the Sobells' tables and related discussion do not convey the reality that is evident when the actual incarceration records of each of the controlled-drinking subjects are analyzed individually" (p. 173). In other words, the difference contended by Pendery et al. (1982) was one of interpretation. Since the findings of Pendery et al. (1982) were presented out of context and were incomplete, a review of the full data set is needed in order to determine which interpretation of the data is more appropriate.

Although Pendery et al. (1982) made inferences supposedly based on hospital and jail incarceration data, it is notable that they did not actually present jail incarceration data in their article. We reported, and Pendery et al. (1982) reiterated that we reported, that over the course of the entire 2-yr follow-up period the traditionally-treated Ss spent more time incarcerated in jail than did the experimentally-treated Ss (e.g. for the first 6 months of follow-up, the experimentally-treated Ss spent 1.94% of days in jail, whereas the traditionally-treated Ss spend 9.23% of days in jail). It seems reasonable to conclude, therefore, that Pendery et al. do not take issue with our reports of jail incarcerations. In any event, the arrest records (California Bureau of Criminal Identification and Investigation) for Ss were reviewed in both independent inquiries of the IBTA study.

In Table 2 (p. 172) of their article, Pendery et al. (1982) present accounts from hospital records of rehospitalizations that occurred among experimentally-treated Ss "within approximately 1 year (italics added) of their participation in the research project" (p. 172). Despite the title of the table, it should be noted that two of the rehospitalizations presented in the table for S CD-E 2 in fact occurred after our 2-yr follow-up interval had expired. This is not discernible from Pendery et al.'s (1982) presentation. In terms of hospitalizations that occurred during the first year of follow-up, their table refers to 12 Ss being hospitalized a total of 15 times. Although not presented in the Pendery et al. (1982) article, these hospitalizations account for 9.43% of the days for all Ss in the first year of follow-up. Most of the hospitalizations reported in the first year after discharge occurred at Patton State Hospital, the records of which were available to Pendery et al. The few remaining hospitalizations in their table were reported as follows:

(a) one at another California state hospital (our own follow-up notes reflect that this hospitalization was noted in the Patton record for this S);
(b) one at a Veterans Administration (VA) hospital (the senior author of the Pendery et al. article is affiliated with the VA and has access to such records);
(c) one at a Naval hospital in the San Diego area; and
(d) a brief admission to a county hospital near Patton.

Since Pendery et al. (1982) did not specifically report how their hospitalization data were collected, it is impossible to determine whether any of the hospitalizations were self-reported by the Ss when they were interviewed several years later by Pendery et al. However, unless Pendery et al. (1982) chose not to report complete hospitalization data, which would be unlikely since the effect of the table was to dramatize adverse experiences of the experimentally-treated Ss, it seems probable that Ss were unable to recall other hospitalizations that they had experienced during the first year of follow-up.

With regard to the second year of follow-up, Pendery et al. (1982) presented data on one hospitalization (within Table 2, p. 1972). Since they reiterated our report that by the fourth 6-month follow-up period the experimentally-treated Ss were only hospitalized 0.46% of days and the traditionally-treated Ss were hospitalized 2.39% of days, we assume they do not contest our
data showing that hospitalizations among the experimentally-treated Ss decreased markedly over the course of our 2-yr follow-up.

In our publications, we reported that the experimentally-treated Ss were hospitalized for alcohol-related reasons for a total of 11.34% of all days during the first year of follow-up (Sobell and Sobell, 1973b). In our presentation of data, we used percentage of days hospitalized as a measure because it was compatible with our other measures of drinking disposition, and because we felt it most accurately reflected the actual impact of a hospitalization on an individual’s functioning. For example, in one case a S had one hospitalization that accounted for 0.82% of days in the follow-up interval. Another S also had one hospitalization, but it accounted for 72.68% of all days in the interval. Simply reporting that each S had been hospitalized, would have occluded the important fact that the first S was minimally impacted by the hospitalization, whereas the second S was incapacitated for nearly three-quarters of the interval.

When our records were reviewed in the two independent inquiries, a breakdown of our hospitalization data indicated that in actuality we had recorded 14 of the experimentally-treated Ss as experiencing a total of either 35 or 37 (the two hospitalizations in question accounted for only 2 days, and we reported the higher figure in our published results) hospitalizations during the first year of follow-up. In their article, Pendery et al. (1982) reported that 12 Ss (although 13 Ss are listed in Table 2, only 12 were readmitted during the first year, as 1 S was readmitted 1 yr 19 days after discharge) had been hospitalized a total of 15 times during the first year of follow-up. The detailed examination of our records also revealed that every single hospitalization reported by Pendery et al. (1982) was accurately included in our records. Thus, our published data are based on more than twice as many hospitalizations for these Ss than reported by Pendery et al. (1982). These 22 additional hospitalizations, not reported by Pendery et al., accounted for 1.91% of all days in the first year of follow-up. This reflects the fact that many were of a very short duration. Therefore, with regard to the data base from which we as compared to Pendery et al. (1982) drew inferences, the data sets differed. This is a first, but vital consideration in examining why Pendery et al. (1982) claimed to have reached discrepant conclusions from our own about the hospitalizations. This difference also raises important questions about the Pendery et al. (1982) study. We do not know why Pendery et al. (1982) did not present complete hospitalization data in their article. However, three alternative explanations seem possible:

1. the hospitalization data that Pendery et al. (1982) presented may have included all hospitalizations of which they were aware, and they may merely have failed to calculate a simple cross-check of their data with our own published data which would have revealed that their data were incomplete;
2. Pendery et al. (1982) may have been aware that their data were incomplete but felt that the differences were of little consequence and need not be mentioned in their article; or
3. Pendery et al. (1982) may have gathered more complete hospitalization data than included in their article but for unexplained reasons only presented partial data in their article.

Pendery et al. (1982) drew two conclusions about the Ss’ hospitalizations:

(i) that our reports did not “convey the reality” (p. 173) of the events; and
(ii) that the hospitalizations “characterized these subjects’ continued attempts to practice social drinking” (p. 173).

In examining how well our reports conveyed the reality of Ss’ functioning, it is essential to recall that the IBTA study was comparative by design. Our major measure of outcome was percent of days functioning well, which consisted of combined days of abstinence and controlled drinking. The reciprocal of this was operationally defined as consisting of days of heavy drinking plus days of alcohol-related incarcerations. For example, for the first 6-month follow-up interval (many of Pendery et al.’s data refer to that interval) we reported that the experimentally-treated Ss functioned well for a mean of 68.36% of all days, whereas the traditionally-treated Ss functioned well for a mean of 38.60% of all days. Recalling the analogy to cancer research, functioning well for 68% of the time is hardly an ideal outcome, but it is considerably better than that for the traditionally-treated Ss. In terms of conveying the reality of outcomes, it is readily seen that for the first 6 months of follow-up, we reported the experimentally-treated Ss as dysfunctioning for 31.64% of all days.
Thus, as a group they spent nearly one-third of the interval either drinking heavily, in hospital or jail. The critical feature of our reports was the comparison with the traditionally-treated group, and by avoiding such a comparison Pendery et al. (1982) failed to address the central focus of the IBTA study, that one treatment was significantly better than another.

As we noted over 10 yr ago in our first publication on the IBTA study, the idea that Ss should consider hospitalization as one alternative to either avoid or to curb a problem drinking episode was a component of the treatment that "had frequently been discussed during sessions as an alternative favorable to starting or continuing to drink" (Sobell and Sobell, 1972, p. 48). In this sense, the treatment embodied components similar to those now used in relapse-prevention approaches (Cummings, Gordon and Marlatt, 1980). Interpretations, of course, are hypothetical by nature, but several considerations support the interpretation that we offered. First, the majority of hospitalizations were in fact very short term. Second, Pendery et al. (1982) did not contest our report that hospitalizations markedly decreased for the experimentally-treated Ss over the course of the 2-yr follow-up. This pattern of improvement has also been found for a study of treatment outcome among chronic alcoholics treated by a skills-training relapse-prevention approach (Marlatt, 1983; Chaney, O’Leary and Marlatt, 1978), where the experimental and control groups did not differ shortly after treatment but the experimental group then proceeded to show significant improvement. Marlatt (1983) interpreted these findings as suggesting that:

"Initial high rates of relapse following treatment may not necessarily indicate that a particular treatment program has been ineffective—perhaps the high initial rate of drinking reflects the early high error rate typically found in a learning curve." (p. 1104)

A final point relevant to the interpretation of the hospitalizations, and also to whether the hospitalizations among experimentally-treated Ss reflected their aspirations to be social drinkers, involves a consideration of the entire context of the IBTA study. An examination of the incarceration outcome data for two groups of Ss who were excluded from eligibility for the controlled-drinking goal—the Nondrinker Experimental (ND-E) and Nondrinker Control (ND-C) Ss—reveals exactly the same outcome pattern as that demonstrated by the two controlled-drinking goal-eligible groups. The ND-E Ss received essentially the same experimental treatment as the experimentally-treated (CD-E) Ss, except that it was oriented toward abstinence. Thus, these Ss were also encouraged to intervene early when drinking got out of hand. During the first year of follow-up, the ND-E Ss spent 11.77% of days in hospital and 5.85% of days in jail. This pattern was reversed for the ND-C Ss, who spent 6.29% of days in hospital and 15.38% of days in jail (Sobell and Sobell, 1978). Also consistent with this interpretation, both groups of experimental (CD-E and ND-E) Ss made greater use of outpatient therapeutic supports over the course of follow-up than did their respective traditionally-treated control groups. If the hospitalizations related specifically to attempts at controlled drinking rather than at the prevention or management of relapses, then the abstinence goal (ND-E) experimentally-treated Ss would not have been expected to show a similar pattern of results. Moreover, Pendery et al.’s (1982) focus on first year outcomes neglected the vast improvement shown by experimentally-treated Ss during the second year of follow-up. Such striking improvement was not found for the traditionally-treated Ss. If hospitalizations are to be attributed to the treatment received, then so should the subsequent decrease in hospitalizations during the second year of follow-up.

Summary. Our assessment of the portrayal of hospitalizations presented by Pendery et al. (1982), therefore, reduces to two major conclusions. First, there are no discrepancies between the data, except that the data presented by Pendery et al. (1982) were somewhat incomplete. The differing interpretation of hospitalizations presented by Pendery et al. (1982), besides being based on incomplete data, resulted from (1) highlighting early relapses in a group which improved substantially over time and (2) disregarding the comparative outcome of traditionally-treated Ss (who, for example, reported more than three times as many heavy-drinking days as the experimentally-treated Ss over the 2-yr follow-up interval). Second, the fact that all hospitalizations reported by Pendery et al. (1982), and more, were accurately reported in our original data and our published reports provides strong, and ironic, corroboration of the accuracy and completeness of our follow-up data.
Other purported discrepancies in 2-yr treatment outcome findings

Pendery et al. (1982) claimed to have reached different conclusions from our own about the 2-yr outcomes on the basis of "documentary records (such as records of hospitalizations for alcoholism and arrests for drunk driving) that would confirm or refute the evaluations of the original investigators" (p. 172). They asserted that "these data, supported by affidavits and records of interviews" (p. 172), led them to reach different conclusions. In addressing this issue, it is necessary to examine the evidence for the validity of the data presented both by Pendery et al. (1982) and by ourselves.

The validity of Pendery et al.'s data. Pendery et al.'s (1982) conclusions were confined to the experimentally-treated Ss, although they conceded that the traditionally-treated Ss had "fared badly" (p. 173). The questionable reliability of the hospitalization records used by Pendery et al. (1982) has already been discussed, as has the fact that they presented little arrest data and apparently did not take issue with our reports of jail incarcerations. As noted earlier and as will be further discussed shortly, there appear to be no contradictions between objective treatment outcome events reported by Pendery et al. (1982) and those which we reported, other than in Pendery et al.'s (1982) interpretation of events and their presenting outcomes out of context for a comparative study. Thus, the validity or invalidity of the Ss' affidavits assumes some importance, because these would be the only remaining sources of contradiction. In this regard, recall that Pendery et al. (1982) did not establish contact with the Ss until several years after the IBTA study was completed. They did not report the dates when affidavits were obtained nor the conditions under which Ss were interviewed (it appears that in many cases multiple contacts preceded the preparation of affidavits). As the Dickens Committee noted:

"Affidavits show only the conscientious beliefs of those who make them, of course, and not that those beliefs are actually true" (Dickens et al., 1982, p. 34)

And even Dr. Maltzman has noted this problem:

"Our evidence is not simply based upon the verbal reports of these patients. If that's all we had, it would be a tempest in a teapot" (Challenge to Sobell Work will have Broad Impact Says Review Group Chief, The Journal, Addiction Research Foundation, Toronto, August 1982, pp. 1, 3)

It is also worthy of note that in commenting on the Rand Report in 1976, Dr. Pendery was said to have "called the researchers 'naive' for having relied on 'notoriously unreliable' information from alcoholics" (Rand Study on Alcoholics Draws Storm of Criticism, Los Angeles Times, 12 June, p. 1).

Unfortunately, we have never been given the opportunity by Pendery et al. to examine or respond to the Ss' purported affidavits. However, in the course of the inquiry conducted by the Dickens Committee, the affidavit of 1 S who figured prominently in the conduct of the Pendery et al. (1982) study became available. If this S's affidavit is representative of the care and accuracy with which the experimentally-treated Ss have provided sworn statements about their experiences several years earlier, then the major remaining source of possibly contradictory data presented by Pendery et al. (1982) can be seen as having highly questionable validity. As part of both independent external inquiries into the IBTA study, an abundance of evidence was submitted that cast grave doubts on the veracity of the one available S affidavit. Perhaps the most telling documentation was the presentation of handwritten and dated correspondence received from the S, himself, and a tape-recorded interview, all of which occurred during a period when the S, by sworn statement, denied having had any contact with us whatsoever. In addition, statements by former research and hospital staff, as well as other evidence, contradicted the affidavit. The Dickens Committee concluded:

"More will not be made of this subject's role. Not least because the effort the Committee felt it would have to make to authenticate all of the many materials we were given to refute his sworn and unsworn statements and published newsmedia statements would have been disproportionate to the instruction to be gained from faulting a person struggling for significance. His good faith and conscientiousness are
not in question, but his recall, changing allegiances and self-attested willingness to subordinate his own perception of the truth in order to say what he feels investigators want him to say, cause the Committee to consider his evidence unreliable." (Dickens et al., 1982, pp. 104–105)

The validity of the Sobells' data. In evaluating the validity of our data, it is important to note that we have retained the original data on which our publications were based, and that documentation was carefully scrutinized in the two independent inquiries. Many types of data were gathered. The accuracy of our records of hospitalizations has already been discussed. Other official record data include complete arrest records, complete driver records, detoxification center admission records, death certificates and autopsy reports and a variety of correspondence with collateral agencies. In addition, our data include complete records of follow-up contacts, copies of correspondence sent to Ss and collaterals and the originals of correspondence received from Ss and collaterals (in many cases this includes postmarked envelopes), tape-recorded interviews with Ss (including all experimentally-treated Ss), telephone logs of follow-up contacts and original data summaries such as those for the variable of residential status and stability. The Dickens Committee noted that our records had been "maintained in fastidious detail which at moments bordered on the obsession, such as in the filing of personal Christmas cards and envelopes" (Dickens et al., 1982, p. 110). The Congressional investigator further noted that:

"The correlation between your notes of contacts with patients, your phone logs and the tape recordings of those contacts have convinced me that your report of your study was made in good faith." (James E. Jensen, personal communication, 23 March 1983).

Several aspects of our follow-up records support the validity of our data. First, in addition to our records accurately including all events (almost all adverse in nature) found by Pendery et al. (1982) in official records, several other events mentioned by Pendery et al. are also accurately noted in our records. These include reports by Ss of their use of therapeutic supports (such as Alcoholics Anonymous and Antabuse), reports of unemployment and physical disability and so on. For example, Pendery et al. (1982) noted that S CD-E 17 had a spinal problem which necessitated surgery, and that "after he was discharged" (italics added) from the research project, his drinking worsened and he lost his job" (p. 173). Both his spinal problem and his unemployment are noted in our records, although Pendery et al. (1982) failed to note that his job loss occurred approx. 18 months after his participation in the study. For another S (CD-E 9) Pendery et al. (1982) reported that he had been treated for Parkinson's disease. They stated that the S had actually pretended to have the disorder as a guise for obtaining Valium and other medications. Whether or not this S truly had Parkinsonian symptoms is not known to us and is irrelevant to the present point. What is relevant, however, is that although there was no need to mention it specifically in our publications (the S is listed, however, as disabled in our published data summaries regarding employment status), our records show that both the S and a collateral reported early in follow-up that he was being assessed for a neurological problem and that the S, himself, reported the use of prescribed Valium shortly after follow-up had commenced. Apparently the neurological indications were sufficient for him to have been medically diagnosed as having Parkinson's disease and his having been prescribed Artane, a drug specifically used for treating that disorder. Thus, even in terms of fine details, Pendery et al. (1982) reported no specific events that were not contained in our follow-up records gathered several years earlier. Hence, the first point favoring the validity of our data over those of Pendery et al. (1982) is this very redundancy—that Pendery et al. reported no specific events which were not also documented in our records.

With regard to the possibility of experimenter bias having influenced the reports given to us by Ss, we cannot rule out such effects, as we stated in 1973(b). However, it should be noted that although it was not mentioned in our publications, it became apparent during the course of the Dickens Committee's investigation that it was actually the case that more than half of the follow-up interviews conducted during the first year of follow-up were performed by research staff other than ourselves, and those individuals have attested to that fact. It is also very important to recognize that the Ss validly reported to us what Pendery et al. (1982) considered to be highly adverse events.
Of further interest is the fact that the experimentally-treated Ss differed from the traditionally-treated Ss in outcome in terms of reports of abstinence as well as reports of limited drinking. More specifically, for the first year of follow-up the experimentally-treated Ss reported being abstinent for a mean of 45.29% of all days as compared to 25.66% of all days for the traditionally-treated Ss. For the second year of follow-up, the experimentally-treated Ss reported being abstinent for a mean of 62.60% of all days as compared to 36.46% of all days for the traditionally-treated Ss. If experimenter effects had occurred, one would have expected the major difference between groups to be in terms of reports of limited drinking. Such a difference was found (about 15% of days on the average), but it was much less striking than either the difference in abstinence days or the previously mentioned difference in days of heavy drinking. Moreover, the findings of Pendery et al. (1982) would, of course, be at least as likely to be vulnerable to experimenter effects as our own, in addition to distortions produced by the passage of time and problems of recall.

The specificity of the data we gathered, and the many validity checks we performed, are also relevant to conclusions about the validity of our data. As reported by the Dickens Committee, the data we gathered were highly specific (sometimes even including the brands of alcoholic beverages). The Dickens Committee also noted (Dickens et al., 1982, pp. 88-89) numerous instances in our follow-up records of our reconciling conflicting reports, as we had previously reported in the literature (Sobell and Sobell, 1973b, 1976). In most cases this resulted in accepting the greater amount of consumption reported from the conflicting sources. We also explored the validity of our Ss' pretreatment self-reports of arrests; those analyses constituted one of the first such studies to appear in the literature (Sobell, Sobell and Samuels, 1974). Another important point is that our data were gathered with few exceptions (as noted elsewhere, e.g. Sobell and Sobell, 1976) prior to our even having become aware that Pendery et al. had an interest in our work. For example, our 1-yr treatment outcome report had already been accepted for publication before we knew of their interest.

An issue which needs some explication is the frequency with which follow-up interviews actually occurred. In all of our publications, we reported that Ss and collaterals had been contacted every 3-4 weeks during the 2-yr follow-up interval, even though we also reported difficulty in locating some Ss for interviews and that the gathering of data in some cases was only completed several months after the end of the 2-yr period. This, of course, was a logical contradiction. But the fact remains that our statements about monthly contacts were incorrect. We were unaware of this until the Dickens Committee actually undertook to log all contacts reported in our records. We had never conducted such a count, and it is of some importance to note why. There are two major reasons. First, the study procedures involved continually seeking information on 70 Ss as well as multiple collateral information sources for each S. In one case, for example, the Dickens Committee noted more than 100 documented attempts to locate a S. This voluminous task, in the context of all 70 Ss in the study, led us to misperceive, some time later when we drafted reports of the study, the frequency with which contacts were achieved. Second, because of the technique used for eliciting data, we were not greatly concerned about how the frequency of contacts might influence our findings, so long as the contacts were multiple.

The above factors, of course, do not justify our error. In our publications, we should have stated that follow-up contacts were 'scheduled' to occur monthly, and that they actually occurred as often as possible. Ideally we should have reported the actual number of contacts. When the Dickens Committee reviewed, entry by entry, all of the information in our follow-up records for the experimentally-treated Ss, it was found that the Ss had actually been contacted an average (mean) of 14.90 times over the course of follow-up and that collaterals had actually been contacted an average of 20.95 times. We should note that the criteria used by the Committee for determining that a contact had occurred were somewhat conservative. For example, when a S was incarcerated for longer than a month and the agency was repeatedly contacted to verify the S's presence, the Committee considered that only one contact had occurred. Nevertheless, the discrepancy in frequency of contacts was readily apparent and thus raises the question of how data for intervals longer than 1 month were gathered.

The technique we used for gathering drinking disposition data was basically the procedure which we later labeled as a timeline follow-back technique. This technique involves several strategies [such as the use of anchor or landmark events and bracketing, see Sobell, Maisto, Sobell, Cooper. Cooper
and Sanders (1980) for a description) to aid Ss in recalling specific data. The technique was later demonstrated in studies, begun in 1975, to have good reliability and validity, even for retrospective periods as long as 1 yr (Sobell, Maisto, Sobell and Cooper, 1979; Maisto, Sobell, Cooper and Sobell, 1979, Cooper, Sobell, Maisto and Sobell, 1980). Certain aspects of the technique, such as the use of anchors or landmark events, have recently received support in the area of memory research for enhancing the accuracy of retrospective recall (Loftus and Marburger, 1983).

The question that might still be raised, however, is whether these techniques were actually developed and used in the IBTA study. In this regard we can offer the following evidence. The basic timeline technique, although not yet given that title, was used in a study we conducted at the Orange County Department of Mental Health Alcoholism Services in 1972 (at the same time when we were still completing follow-up for the IBTA study). In that study, daily drinking disposition data were gathered for a 6-month posttreatment retrospective period from alcoholic Ss who had been treated at private alcohol recovery hospitals and from their collateral informants. The data from that study were not published until 1979 (Maisto, Sobell and Sobell, 1979), although we referred to the study in 1976 (Sobell and Sobell, 1976). It also remains a fact that the actual frequency of contact with Ss and collaterals in the IBTA study is among the highest ever reported in the alcohol treatment research literature. For example, in a recent review of 35 alcohol treatment outcome studies published from 1976 through early 1980, we found that the mean number of follow-up contacts reported was 5.9 (Sobell and Sobell, 1982). We believe that taken in concert with all of the other evidence supporting the validity of our findings, the impact of less frequent contacts with Ss and collaterals than we had reported is negligible, and our findings remain some of the most complete and cross-validated data ever reported in the literature. This is especially the case in view of the two independent external inquiries supporting the veracity of our data.

Conclusions regarding the 2-yr follow-up

The preceding discussion supports several conclusions. Pendery et al. (1982) did not take issue with our data reporting improvement for Ss over the course of follow-up, and they reported no verifiable events which were also not a part of our follow-up records. A large amount of documentation supports our published follow-up reports, and cross-checks conducted in the two independent inquiries corroborated those records. Thus, other than the possible contention by Ss (several years after the fact and after having apparently been contacted on multiple occasions by Pendery et al.) that our reports are inaccurate, what discrepancies really exist? Is the conflict just a "tempest in a teapot?" It is a reasonable assumption that the data published by Pendery et al. (1982) represent their strongest evidence of discrepancies with our reports. In addition, they had available our published reports, including detailed individual outcome data, to use in conducting and reporting their study. The major points of contention appear to be our assertion that "many of the CD-E subjects engaged in limited, nonproblem drinking throughout the follow-up period" (Sobell and Sobell, 1978, p. 155), and as Pendery et al. (1982) put it:

"their (the Sobells) conclusion that training directed toward controlled drinking is an effective therapy for gamma alcoholism." (p. 174)

In addressing the above issues, we must again refer to the dangers of taking statements out of context. Pendery et al. (1982) could have chosen, for instance, more appropriate passages for citation, such as passages which we, ourselves, emphasized. If the reader carefully reviews our publications, it will be readily apparent that we did not claim the experimentally-treated Ss to have ideal, or drunkenness-free outcomes. All of the experimentally-treated Ss were reported to have had some days of heavy drinking over the course of follow-up, and adverse consequences they experienced were accurately reported in our publications. Many Ss also reported a substantial number of days of limited drinking not followed by adverse consequences. Pendery et al. (1982) have not presented any evidence contradicting those reports.

With regard to the efficacy of controlled drinking as a treatment goal, we were quite cautious and, as mentioned earlier, were careful to state that the design of the study precluded drawing conclusions separating that goal from the method of treatment. It seemed quite possible to us, and remains so, that the better outcome for the experimentally-treated Ss primarily reflected the influence of the broad-spectrum behavioral treatment, and perhaps also the nonefficacy of the
traditionally-oriented treatment for this population. To use a quote from the same page as cited by Pendery et al. (1982):

"Given these results, one might conclude that the treatment goal of controlled drinking contributed more to successful outcomes than did the method of Individualized Behavior Therapy. However, such a conclusion cannot be adduced from the present data. The results of the IBT study are not only complex but unique, making interpretation and comparison of those data similarly complex, especially with respect to drinking outcomes." (Sobell and Sobell, 1978, p. 155; italics in original)

We have always recognized and been careful to draw attention to the fact that interpretations of findings may differ among individuals. In fact, in our publications we presented detailed outcome data for both groups and individual Ss in order that readers would be free to make their own interpretations. Pendery et al. (1982) were wrong in their attribution to us of the conclusion that controlled drinking was an effective treatment for yummia alcoholics. What the IBTA study showed was that the experimental treatment was a more effective treatment than the intervention based on conventional wisdom which was otherwise available for alcoholic patients in treatment at Patton State Hospital. That finding remains valid, as does the IBTA study.

**Long-term Treatment Outcomes**

The third major focus of the Pendery et al. (1982) critique was the presentation of long-term outcome data for only the experimentally-treated Ss. From the foregoing, it should already be clear that there are great limitations to the knowledge that can be gained from presenting findings for only one group in a comparative study. Nevertheless, even within this constraint, an evaluation of the long-term outcomes may be instructive.

Pendery et al. (1982) discussed treatment outcomes through the end of 1981. Thus, the follow-up period for individual Ss varied from approx. 10½ to 11½ yr. It is notable that long-term treatment outcome information of any sort is a rarity in the alcohol literature. For example, Vaillant (1980, 1983) summarized follow-up data from six long-term studies, of which only four were treatment evaluations, and stated that:

"it is unfortunate that there are no more careful long-term (more than six years of follow-up) studies of alcoholism in the world literature." (Vaillant, 1980, p. 18; italics in original)

Despite being an outspoken defender of conventional alcoholism treatment methods, after reviewing the results of an 8-yr follow-up of a Massachusetts hospital-based treatment program with which he was affiliated, Vaillant (1983) was forced to conclude that:

"after initial discharge, only 5 patients (4.7%) in the Clinic sample never relapsed to alcoholic drinking, and there is compelling evidence that the results of our treatment were no better than the natural history of the disease." (p. 284, italics in original)

Comparing those results with the other five long-term studies, Vaillant (1983) further concluded:

"Once again, our results were no better than the natural history of the disorder." (p. 285).

As reflected in the following examples, even short-term positive outcomes are a rarity, especially for chronic alcoholics treated in public programs. In a recent article, Gottheil, Thornton, Skoloda and Alterman (1982) reviewed the results of three studies conducted in public programs and found 6-month abstinence rates of only 18–19%. Similarly, Dwoskin, Gordis and Dorph (1979) studied the fates of 183 alcoholic halfway house clients after discharge and found that only 19% of the Ss reported they had maintained abstinence for as long as 8 months postdischarge. Further, in an evaluation of 781 clients treated at public programs in the U.S.A., Polich et al. (1981) found that only 7% of the Ss claimed to have maintained continuous abstinence over the course of 4 yr. Finally, Gordis, Dorph, Sepe and Smith (1981) found that only 9% of 5578 persons treated for alcoholism at a hospital-based program in New York City had maintained abstinence for as long as 1 yr following treatment. They concluded that "contemporary alcoholism treatment is, at best,
of limited effectiveness" (Gordis et al., 1981, p. 509). They also went on to recommend that "potent new therapy is needed, and only further research can provide it" (Gordis et al., 1981, p. 521).

As Costello et al. (1977), among others, have noted, "the studies reporting higher success rates are more likely to utilize a management strategy of poor-prognosis subject exclusion" (p. 314). Using a criterion of success which did not require complete abstinence, Costello et al. (1977) reviewed the results of 80 studies published from 1951 to 1975 and concluded that "at 1-yr follow-up ... programs with limited resources and poor-prognosis patients ... can expect a success rate of only 15%" (p. 311). Even for good-prognosis patients, long-term outcomes are seldom impressive. For example, in a 4-yr follow-up of patients treated at a private treatment program, Pettinati, Sugarman, DiDonato and Maurer (1982) found that only 22% of a sample of 200 patients had maintained abstinence with good adjustment for 4 yr following treatment. Thus, with a poor-prognosis sample, such as that which characterized the Ss in the IBTA study, one would not expect to find exceedingly positive outcomes over an 11-yr period associated with any method of treatment.

In recent years, the impact of environmental factors on treatment outcomes has become a topic of major interest. In this regard, Moos and Finney (1982) have noted the influence of posttreatment events on treatment outcomes and have pointed out the logical difficulty in making attributions of long-term treatment effectiveness:

"As a case in point, six hours of outpatient treatment may have some short-term benefit for a client, but, since any such benefit is likely to be 'diluted' by clients' stressful life situations, there is little reason to expect any substantial effects four years after treatment. It makes even less sense to expect strong evidence of treatment benefits ten years after treatment. These considerations highlight the need for a paradigm shift in evaluations of alcoholism programs." (Moos and Finney, 1982)

Returning to the issue at hand, the long-term outcome data presented by Pendery et al. (1982) were, with the exception of mortality findings, basically global reports of various types of events, many of which occurred several years after treatment (e.g. 1979-1981). Presented in isolation, these findings shed no light on the research question which the IBTA study was designed to answer—whether the experimentally-treated Ss had superior outcomes to the traditionally-treated Ss. Since we have gathered no outcome data of a similar nature for the traditionally-treated Ss beyond the initial 2-yr period of study and know only that Pendery et al. (1982) reported that those Ss "fared badly" (p. 173), we can offer little with regard to general outcomes other than to point out that the information presented by Pendery et al. lacks meaning for evaluating the IBTA study.

One aspect of the long-term outcome data presented by Pendery et al. (1982), however, are subject to comparative analysis. These are the findings on mortality.

Mortality findings

Before discussing the comparative mortality data from the IBTA study, it will be of value to provide some background for the interpretation of alcohol-related mortalities. This is important, because even without knowledge of the mortality rate for the traditionally-treated Ss, a review of the mortality literature shows that the rate reported by Pendery et al. (1982) for the experimentally-treated Ss, especially through the first 6 yr following treatment, was below the expected rate of mortality for such a population. However, Pendery et al. (1982) did not discuss the obtained mortality rate for the experimentally-treated Ss in relation to the vast literature on alcohol-related mortality.

Before proceeding, some mention of the findings of long-term follow-up studies of mortalities among treated individuals is relevant to the discussion that follows. In summarizing the mortality findings of three long-term studies of treated alcoholics, Vaillant (1980) reported that the mortality rates for these three studies were 25% at 7 yr, 14% at 10 yr and 22.5% at 9 yr. In a very recent article, Taylor, Combs-Orme and Taylor (1983) reported that 19.4% of 1289 treated alcoholics died "of alcoholism or its complications" (p. 17) over a 5- to 8-yr posttreatment interval, and further, that more than one-quarter of those deaths occurred within the first posttreatment year. Fitzgerald, Pasewark and Clark (1971) reported a mortality rate for state hospital-treated alcoholics of 12.8% over 4 yr. Tashiro and Lipscomb (1963), in a study of alcoholic mortality that included patients
treated in California state hospital programs, found that 7.3% had died within a 2- to 5-yr period following treatment. For alcoholics treated in public programs, Polich et al. (1981) reported a 14.5% mortality rate over 4 yr. Finally, Pokorny, Miller and Cleveland (1968) found that 4.5% of alcoholics treated at a Veterans Administration hospital died within 1 yr of treatment discharge. The above studies notwithstanding, there are also several summaries of similar mortality data readily available in the literature (e.g. Schmidt, 1980; Costello, Parsons-Manders and Schneider, 1978).

Even with good-prognosis patients mortality rates are still relatively high. Ōjesjö (1981), for example, in a 15-yr longitudinal study where a majority of the Ss were not physically dependent on alcohol, found a 26.4% mortality rate. Pettinati et al. (1982) found that 7% of alcoholics treated at a private program died within a 4-yr follow-up interval. On occasion, the mortality rates associated with treatment of good-prognosis patients have been found to be surprisingly high. For example, a study conducted in 1974 at the Donwood Institute (Toronto) evaluated mortality over 5 yr for a cohort of 154 patients treated for alcoholism. It should also be pointed out that patients with serious organic complications were deliberately excluded from this study, and the sample was described as mostly "married, high school or college graduates, and in professional, semiprofessional, managerial or proprietary occupations" (DeLint and Levinson, 1975, p. 386). Thus, these patients were high in social stability and personal resources in comparison with the patients in the IBTA study. Yet, within 5 yr after treatment, 22 patients (14.3%) had died. The authors concluded that:

"In our view the majority of deaths in the Donwood sample can indeed be linked to alcohol use and some related factors." (DeLint and Levinson, 1975, p. 387)

Clearly, it would be absurd to attribute the high mortality rate in the Donwood sample to the treatment they had received 5 yr earlier.

The foregoing provides an appropriate context for considering the mortality findings presented by Pendery et al. (1982). Pendery et al. (1982) contended that 4 of the experimentally-treated Ss "eventually died alcohol-related deaths" (p. 174). The circumstances of death and whether all of the deaths can reasonably be considered alcohol related will be discussed shortly. First, however, it is important to note the way in which Pendery et al. (1982) reported the deaths of the experimentally-treated Ss. The dates of death only appeared among the footnotes at the end of their article, and even then only in parentheses as the dates of official records. Thus readers would not readily have discerned that the deaths actually occurred 6, 81, 101 and 11 yr following the Ss' participation in the IBTA study. That is, *no experimentally-treated Ss died within the first 6 yr following discharge*, a finding that stands in marked contrast to the expected mortality rate for this sample. Pendery et al. (1982) did not present comparable mortality data for the traditionally-treated Ss, although they did state in a footnote regarding that group that "Eleven subjects and the widows of two others were interviewed. Another had already been reported to have died ..." (p. 175, footnote 23). Such a statement clearly implies that only three of the traditionally-treated Ss were known to have died.

At the request of the Dickens Committee, we sought mortality information for Ss in both groups. This involved obtaining death certificates, autopsy reports, when available, and other available information. The findings reported by Pendery et al. (1982) for the experimentally-treated Ss were thereby confirmed. However, we found Pendery et al.'s (1982) mention of deaths among the traditionally-treated Ss to be incomplete. Since death records are public records, we simply wrote to the California Bureau of Vital Statistics and asked for copies of any death records on file for the traditionally-treated Ss. Thus, we report here only on deaths known to authorities in the state of California (out-of-state checks were also made on 2 Ss, but proved negative). As a result of this inquiry, we learned that 6 of the traditionally-treated Ss had died through 1981. We then proceeded to gather relevant public records (autopsy and coroners' reports) for the cases so identified. Since these records were easily obtained, and since Pendery et al. (1982) presented outcome data through 1981, it is curious that three deaths among the traditionally-treated Ss were not reported by Pendery et al. (1982), even though the most recent death in the experimentally-treated group had occurred after the most recent death in the traditionally-treated group. A consideration of the circumstances, nature and dates of deaths reveals some interesting comparisons between the
groups. In our view, only three of the four deaths among experimentally-treated Ss can be considered alcohol related. Below, the deaths are discussed on a case-by-case basis. Experimentally-treated Ss are denoted as CD-E and traditionally-treated Ss as CD-C, representing their original treatment group designations.

**CD-E 12.** This S died approx. 6½ yr following his participation in the IBTA study. As reported by Pendery et al. (1982), he died of a massive myocardial infarction. No autopsy was conducted. When we contacted the attending physician and a collateral information source, we confirmed, as reported by Pendery et al. (1982), that the S had been abstinent for about 1 yr prior to his fatal heart attack. However, whereas the attending physician was reported by Pendery et al. (1982) to have stated that the S’s “abstinence was instituted too late to prevent his untimely death” (p. 174), the same attending physician told us that the cardiologist on the case “would not commit himself” that the death was alcohol related (D. F. Rodwell M.D., personal communication, 10 June 1982).

In any event, the opinion of the attending physician and cardiologist are not actually of great consequence in this case. As noted by Knott and Beard (1982), discerning the relationship of alcohol ingestion to cardiovascular consequences in any individual case is a complex task, since other factors (e.g. heavy smoking, nutritional–metabolic abnormalities, life-style patterns) “certainly contribute to the enhanced circulatory disease mortality and morbidity in heavy drinking and alcoholic persons” [p. 333; see also Regan, Ahmed and Ettinger (1981)]. Further complicating the problem of interpretation is evidence that cardiomyopathy (also present in this S) diminishes with abstinence (Baudet, Rigaud, Rocha, Bardet and Boudriarias, 1979; Gunnar, Demakis, Rahimtoola, Sinnio and Tobin, 1975). The direct relationship, if any, between this S’s drinking and his death, therefore, is at least tenuous, as indicated by the absence of mention of an alcohol-related contributing cause of death on his death certificate.

**CD-E 16.** This S, who died 8½ yr after his IBTAstudy participation, was reported by Pendery et al. (1982) to have committed suicide when he jumped off a pier with a blood alcohol level of 0.30%. This circumstance is correct, and the death was assuredly alcohol related. The coroner’s report (K. R. Bell, County of San Diego, Calif., 1979) further indicated that approx. 2 yr before the S’s suicide, he had one leg amputated below the knee. “After that time, he became progressively despondent due to his inability to ambulate normally” (coroner’s report, 1979). He was reported to have attempted suicide approx. 1 yr prior to his death and again 2 weeks prior to his death. After the latter attempt, he received inpatient psychiatric counseling. He left the hospital against medical advice a few days prior to his successful suicide. Our presentation of these details is not to obscure the obvious relationship between his intoxication and his death, but simply to point out that there were concurrent emotional problems which were not present until several years after the S had been discharged from the IBTA study.

**CD-E 14.** This S died 10½ yr following his participation in the IBTA study. The S was reported by Pendery et al. (1982) as having died of respiratory failure due to ethanol intoxication, consistent with the autopsy report. Of some interest, the S was reported (C. P. Speth, autopsy protocol, County of San Bernardino, Calif., 1981) to have left an alcoholic rehabilitation program 5 days before his death, and to have been treated for ‘alcoholic binges’ at a local hospital two other times within 2 days prior to his death.

**CD-E 6.** This S died 11 yr following his participation in the IBTA study. As accurately reported by Pendery et al. (1982), the S died of drowning, with a blood alcohol level of 0.30% (A. D. Mahoney, autopsy protocol, County of San Bernardino, Calif., 1981). The autopsy report also noted that the S had “walked away from an alcoholic rehabilitation center” 4 days prior to his death.

Thus, 4 of the experimentally-treated Ss died within 6½–11 yr following the IBTA study, 3 for clearly alcohol-related reasons. It should be noted, for comparative purposes, that the alcohol-related deaths were in all cases the result of accidental or intentional occurrences while acutely intoxicated, rather than from chronic diseases well known to be related to alcoholism and more indicative of chronic heavy drinking. Upon consideration of the deaths among experimentally-treated Ss, a further observation of interest can be made. If all 20 Ss in the experimentally-treated group are ranked according to the number of days we reported them to have functioned well during the initial 2 yr of follow-up, it can be shown that the 4 Ss who later died ranked 15th, 16th, 17th and 20th in terms of functioning (i.e. with 20th the poorest outcome). Moreover, in terms of dates
of death, their order of death is completely predictable from our follow-up data presented several years earlier. With no exceptions, the worse a S was reported to have fared over the 2-yr of follow-up, the earlier the S died. In this morbid sense, the data for the experimentally-treated Ss might be said to have predictive validity, and it would be of interest to know whether other studies would show similar relationships. The deaths among traditionally-treated Ss will now be considered in similar detail.

CD-C 19. This S died 2.5 yr following his participation in the IBTA study. His death was previously reported (Sobell and Sobell, 1976), and resulted from thermal burns with a blood alcohol level of 0.28% (E. Carpenter, autopsy report, County of Los Angeles, Calif., 1974). The S had apparently been smoking in bed in a boarding house when the fire occurred.

CD-C 9. This S died 5.1 yr following his participation in the IBTA study. The S's death certificate indicated that he had died of "acute pulmonary embolia" due to "chronic obstructive pulmonary disease". Although an autopsy was performed, we were unable to gain access to the autopsy report. Since alcoholics have a proneness for respiratory disease (Lyons, 1982), interpretations of the possible contribution of alcohol ingestion to death from pulmonary disease are difficult without additional information (for the same reasons cited in the interpretation of alcohol involvement in the death of S CD-E 12). Thus, it is presently unknown whether this S's death was alcohol related.

CD-C 14. This S died 8.5 yr following his participation in the IBTA study. The death certificate indicated the immediate cause of death as "coronary arteriosclerosis". Although the autopsy protocol (I. Root, County of San Bernardino, Calif., 1979) indicated that no blood alcohol level was found, the autopsy report stated:

"He apparently (sic) been drinking heavily lately. There was (sic) bottles of Brandy found. Some full, some empty".

Therefore, it is possible, but not definite, that this S's death should be considered alcohol related.

CD-C 13. This S died 9.1 yr following his participation in the IBTA study. His death certificate listed "respiratory arrest" due to "chronic obstructive pulmonary disease, bronchopulmonary pneumonia" as the immediate cause of death, and "alcoholic liver disease" as a contributing cause of death. The autopsy report (E. O. Johanson, autopsy protocol, County of Riverside, Calif., 1980) indicated that when the S was admitted to the hospital for treatment of pneumonia, "he had symptoms of alcohol withdrawal". This S's death, therefore, was alcohol related.

CD-C 10. This S died 10 yr following his participation in the IBTA study. Although we were unable to obtain the autopsy report, his death certificate indicated the immediate cause of death as "bronchial emphysema, cigarette smoking", with "chronic liver disease (alcoholic)" as a contributing cause of death. Thus, there is little doubt that this S's death was alcohol related.

CD-C 12. This S died 10.1 yr following his participation in the IBTA study. His death certificate and autopsy report (D. M. Satsuyama, County of San Diego, Calif., 1981) listed the immediate cause of death as "cardiorespiratory failure" due to "hypoglycemia; cirrhosis, liver, Laennec's". The coroner's report (R. S. Grubb, County of San Diego, Calif., 1981) indicated that the decedent "had been drinking a great deal during the day and evening" of the day before he was found dead. Laennec's cirrhosis is a synonym for "alcoholic cirrhosis" (Harvey, Johns, Owens and Ross, 1976). Thus, this S's death was alcohol related.

As described above, 6 of the traditionally-treated Ss died through 1981, and two of the deaths occurred before any deaths occurred among experimentally-treated Ss. Four of the six deaths were clearly alcohol related. In the fifth case, the S was reported to have been drinking heavily prior to his death, but it is unclear whether alcohol ingestion was a contributing cause of death. In the remaining case, information was not available to make a determination of alcohol involvement.

With the exception of the one alcohol-related accidental death, the causes of death among traditionally-treated Ss were chronic diseases frequently (directly or indirectly) related to chronic alcoholism. This is consistent with the Ss reports several years earlier of substantial amounts of very heavy drinking. Unlike the experimentally-treated Ss, however, for the traditionally-treated Ss the dates of deaths bore no systematic relationship to the Ss' status over the 2 yr of follow-up in the IBTA study. In fact, four of the six deaths occurred for Ss who over the earlier course of
follow-up ranked within the top seven in their group in terms of percentage of days functioning well.

Although we place little credence on attempts to relate a long-term outcomes to a treatment received 11 yr earlier, if one chooses to draw such inferences, the compelling conclusion is that even by the ultimate criterion of outcome emphasized by Pendery et al. (1982), alcohol-related mortalities, the experimentally-treated Ss fared better over the long run than did the traditionally-treated Ss. We are left again with a distinct reaffirmation of the IBTA study results. Taken in context, then, what was the point of the Pendery et al. (1982) critique? One can only present interpretations, and in our concluding section which follows, we present our interpretation of what has occurred.

**Scientific Revolution in the Alcohol Field**

We have already suggested that Kuhn's (1970) well-known theory of scientific revolutions can provide a meaningful framework for interpreting recent events in the alcohol field. Since not all readers will be familiar with Kuhn's theory, however, we will first briefly outline his major tenets. As we are not philosophers of science, we make no claim to technical accuracy. Thus, what follows is best considered as our interpretation of Kuhn's theory.

A central proposition argued by Kuhn (1970) is that major changes in scientific thought do not occur gradually as the result of assimilating new knowledge, but rather abruptly, in a revolutionary manner. The conflict involves advocates of an older, established view, and those advocating its replacement by a new and different set of ideas. In describing what a particular scientific view is, Kuhn invokes the concept of scientific paradigms. From a psychological perspective, a paradigm can be viewed as a cognitive set. It is broader than a theory. It is a shared belief of a scientific community. It represents a way of thinking about a problem, including specifying what research questions are worth asking, and how to go about seeking answers. Scientific revolution describes the process of paradigm change.

As shown in Fig. 1, any field begins with a preparadigmatic period, during which there are competing schools of thought. Eventually a prevailing view emerges, and Kuhn calls this dominant paradigm 'normal science'. Normal science is the textbook science, the most commonly accepted

---

Fig. 1. The process of scientific revolutions, derived from Kuhn (1970).
explanation of a phenomenon. Moreover, most scientific activity occurs within the context of normal science, where scientists of like view seek to answer research questions within the context of an accepted paradigm. An unusual feature of the alcohol field is that the present normal science did not have an empirical basis. Basically, it derived from lay theories (most notably the view expressed in the book Alcoholics Anonymous, 1939) merged with scientific speculation (most notably Jellinek's 1960 book, The Disease Concept of Alcoholism). These traditional concepts have been summarized by Pattison et al. (1977) and Heather and Robertson (1981) and will not be elaborated on here. The important point is that the traditions to be defended, so-called conventional wisdom, did not have their roots in scientific research.

The trigger to scientific revolution lies in the accumulation of research anomalies—findings that are inconsistent with the normal science. Since our knowledge is seldom complete, it is to be expected that research will yield anomalies. Some amount of anomalies must be tolerated. Some will be spurious findings that cannot be replicated. Others will be erroneous, the result of flawed research. And the few that are neither spurious nor results of flawed research can usually be ignored as findings that are likely to be explained over time as the state of knowledge improves. However, when anomalies are persistent, and when the set of anomalies builds to a point where they can no longer be ignored, a crisis ensues for normal science. When this happens, new persons entering the field, lacking commitment to the normal science, are struck that the prevailing view cannot adequately account for the anomalies. Thus, they, and others not so strongly committed to the dominant view, begin to search for better explanations.

Eventually, alternative paradigms are proposed which seem to better account for the evidence. Such paradigms gain adherents, who see them as providing a better guide for research. As the new ideas gain stature, and particularly as they become cited in textbooks, the normal science is threatened and an intense polarization develops. Kuhn (1970) likens this process to a political revolution, where there are competing camps, each committed to a different and incompatible set of views. Examples of attempts to develop alternative paradigms in the alcohol field can be found in Pattison et al. (1977), Beauchamp (1980) and Heather and Roberston (1981). It should be noted that besides the issue of reversibility, there are a great many areas (e.g. progressivity of the disorder) where research findings diverge from the traditional view, and that it is the failure of the normal science in the alcohol field to explain these diverse anomalies, as well as the failure to present a research foundation for traditional views, which has led some to seek alternative paradigms.

Once competing paradigms are introduced, there follows a revolutionary battle, with proponents of each paradigm arguing in its own defense and often seeking to discredit the other view. At this point it becomes most clear, as Kuhn contends, that science is a social process. Kuhn (1970) noted that "the parties to a revolutionary conflict must finally resort to the techniques of mass persuasion, often including force" (p. 93). We contend that it is this stage of revolutionary battle which characterizes the alcohol field today. Kuhn (1970) also noted that:

"The issue is which paradigm should in the future guide research on problems many of which neither competitor can yet claim to resolve completely. A decision between alternative ways of practicing science is called for, and in the circumstances that decision must be based less on past achievement and more on future promise... A decision of that kind can only be made on faith." (pp. 157–158)

In general, the revolutionary battle may be followed by any of three outcomes. The alternative view may be shown to be a less adequate explanation than the view it is proposed to replace. Or, despite its better explanatory value, the alternative view might be effectively suppressed. Or, the new paradigm might prevail. When the latter outcome occurs, the revolutionary science of today becomes the normal science of tomorrow, only to fall prey at some later date to the same revolutionary process.

Another important aspect of the scientific revolution currently underway in the alcohol field concerns the locus of battle. In the physical sciences, often only scientists can understand why the issues in dispute are important, and the revolutionary battle takes place within the scientific community. Thus, Kuhn (1970) stated:

"One of the strongest, if still unwritten, rules of scientific life is the prohibition of appeals to heads of state or to the populace at large in matters scientific" (p. 168).
In the alcohol field, however, the normal science view is, in one way or another, common knowledge among the general public, as is the controversy over controlled-drinking research.

As already mentioned, the research literature in the alcohol field is fraught with multiple areas of important anomalies that cannot be explained by the traditional view. Of these anomalies, the issue of reversibility has been the most publicly prominent and the most vigorously debated. Heather and Robertson (1981) have suggested that this is because:

"the belief in the therapeutic necessity for total and lifelong abstinence, a necessity which logically entails the postulation of an irreversible deficit. is the belief which unites, more than any other, proponents of disease conceptions of alcoholism. It is the nature of this dominant treatment goal and its transformation into a widely promulgated 'abstinence ideology' which has had the greatest possible influence on the ways in which alcoholism is popularly viewed in our society." (p. 246)

The controversy in the alcohol field regarding the issue of reversibility also exemplifies other aspects of the process of scientific revolution as described by Kuhn (1970). Undeniably, there exists an intense polarization of views. There have been distinct attempts at mass persuasion, particularly via the media, as well as attempts to personally discredit advocates of the competing view. There have also been attempts to suppress the publication of anomalous findings. Selzer (see Davies, 1963), commenting on Davies' (1962) controversial report of successful nonabstinent outcomes, noted that he had found similar findings in a previous study and been encouraged by those who had funded his research to omit mention of those "embarrassing" findings in his report. The best-known attempt at suppression concerned publication of the first Rand Report, where efforts were made to delay or modify publication of those findings. Roizen (1977), for instance, commented that

"It is a matter of public record that both officials at Rand and a United States Senator were asked to delay publication of the report until its contents could be reviewed and its data, perhaps, reanalyzed." (p. 171).

Armor et al. (1978), the authors of the Rand Report, noted that the National Council on Alcoholism (NCA), following release of the report, held a press conference (with Dr Pendery as a major spokesperson) where "the Rand Report was characterized as 'dangerous' and 'unscientific.' despite the fact that the NCA had not yet seen the report" (p. 213). Armor et al. (1978) provide a comprehensive and documented summary of this conflict in their book (Appendix B). A further repercussion of the Rand Report was described by Lender and Martin (1982) in their book Drinking in America. They reported that:

“When the Journal of Studies on Alcohol published a number of articles with conclusions similar to Rand's, one major private contributor to alcoholism research proceeded to cut the Rutgers Center of Alcohol Studies, the Journal's publisher, out of his will.” (Lender and Martin, 1982, p. 193)

Kuhn (1970) has also emphasized that textbooks are the purveyors of normal science. in that the information imparted in textbooks, although often in imperfect ways, forms the core knowledge and orientation for new generations of scientists. Thus, in an ideological struggle, such as the scientific revolution we believe is underway in the alcohol field, one should expect that pressure will be brought to bear on what is or is not included in textbooks. The evidence supports this prediction. In a public lecture, tape-recorded for sale, Dr Pendery stated:

“Obviously what we want is a correction in the textbook literature. That's I think, if we have one specific goal, that has to be it.” (Blume, Wallace, Pendery and Tuchfeld, 1983)

In many ways, therefore, the evidence compels the conclusion that the alcohol field exemplifies a scientific revolution in progress. The immediate outcome of the battle is far from assured, because research evidence is not necessarily the basis upon which resolution will occur. There is some basis for asserting, however, that change will eventually come about. Two reasons are foremost. First, the traditional view, or normal science, is in many ways not consistent with the body of research
evidence. Second, and perhaps more important, there is a pronounced lack of evidence that the present normal science has yielded effective methods of treatment. A central impetus for change, therefore, is that the traditional view no longer inspires progress.

Acknowledgements—This article would not be complete if we did not acknowledge the many friends and colleagues whose support helped us through this conflict. To each of you we express our thanks. There are also a few individuals whose help, support and advice made all the difference, and to whom we owe a debt of gratitude. Finally, we are especially indebted to the Addiction Research Foundation for its willingness to support an objective evaluation of the issues.

REFERENCES


