COGNITIVE-BEHAVIOUR THERAPY AND PROBLEM DRINKING:
A META-ANALYSIS

by

Andrew G. Matthew

A thesis submitted in conformity with the requirements
for the degree of Master of Arts
Department of Human Development and Applied Psychology
Ontario Institute for Studies in Education of the
University of Toronto

© Copyright by Andrew G. Matthew 1997
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author’s permission.

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L’auteur conserve la propriété du droit d’auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-51567-2
Acknowledgments

I am deeply grateful to Dr. Barry H. Schneider for his guidance and encouragement in the preparation of this manuscript. I am also grateful to Dr. Anthony Toneatto for his support and helpful advice. Further, I wish to thank John G. Matthew for his thoughtful assistance in editing this manuscript.

Finally, I would like to thank my family and friends for their unending support and reassurance throughout the process of writing this manuscript.
Cognitive-behavioural theories have been proposed to explain why, where, and how much people drink, and why most remain social drinkers while others experience problem drinking. Cognitive-behavioural treatment strategies focus on the cognitively mediated variables believed to be involved in the development and maintenance of problem drinking. The aim of this quantitative review is to determine the overall effectiveness of Cognitive-Behavioural Therapy (CBT) in the treatment of problem drinking according to the literature to date.

An English-language computerised literature search was used to locate studies (n = 12) reporting the results of controlled trials. The results from the original studies were statistically pooled to establish the overall effect of CBT in the treatment of problem drinking. CBT was found to be effective in reducing the alcohol intake of problem drinkers, at least in the short term (mean within-group effect size = .50), but there is little evidence to suggest that it is superior to other forms of treatment. Specific recommendations are made with respect to design methodology in future research.
Table Of Contents

Acknowledgements ...................................... i
Abstract ................................................. ii
List of Tables ........................................... v
List of Figures .......................................... vi

I. Introduction ............................................. 1
   Cognitive-Behavioural Models And
   Strategies Of Problem Drinking ..................... 5
   Efficacy of Cognitive-Behaviour Therapy
   In The Treatment of Problem Drinking ............. 7
   Meta-analysis ......................................... 11

II. Method .................................................. 12
   Literature Search ..................................... 12
   Inclusion Criteria .................................... 13
      Rationale ........................................... 13
   Data Extraction ...................................... 15
   Conventions Adopted ................................ 15
   Methodological Quality ............................... 17
   Statistical Analysis ................................ 17

III. Results ............................................... 20
   Literature Search .................................... 20
   Study Characteristics ................................ 21
   Overall Effect Sizes ................................ 24
   Outlier Effect ....................................... 25
   Overall Effects ...................................... 25
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment Dropout</td>
<td>26</td>
</tr>
<tr>
<td>Different Methods Of Calculating Effect Size</td>
<td>27</td>
</tr>
<tr>
<td>From Non-Significant Results</td>
<td>28</td>
</tr>
<tr>
<td>Methodological Quality</td>
<td>28</td>
</tr>
<tr>
<td>Random Assignment</td>
<td>29</td>
</tr>
<tr>
<td>Integrity Control Of Treatment Delivered</td>
<td>29</td>
</tr>
<tr>
<td>Follow-Up Interval</td>
<td>30</td>
</tr>
<tr>
<td>IV. Discussion</td>
<td>37</td>
</tr>
<tr>
<td>V. Conclusion</td>
<td>37</td>
</tr>
<tr>
<td>References</td>
<td>53</td>
</tr>
<tr>
<td>Appendix A. MeSH Terms Employed In The Computerised Search Strategies</td>
<td>61</td>
</tr>
<tr>
<td>Appendix B. Authors Individually Searched In Medline and Psyclit</td>
<td>62</td>
</tr>
<tr>
<td>Appendix C. Descriptive Data Extracted From Individual Studies</td>
<td>63</td>
</tr>
<tr>
<td>Appendix D. Methodological Quality Scale</td>
<td>66</td>
</tr>
<tr>
<td>Appendix E. References Of Studies That Did Not Meet The Initial Screening Criteria</td>
<td>68</td>
</tr>
<tr>
<td>Appendix F. References Of Studies That Did Not Meet The Final Inclusion Criteria</td>
<td>69</td>
</tr>
<tr>
<td>Table</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>1.</td>
<td>Experimental Characteristics</td>
</tr>
<tr>
<td>2.</td>
<td>Subject Characteristics</td>
</tr>
<tr>
<td>3.</td>
<td>Effect Sizes For Individual Studies</td>
</tr>
<tr>
<td>4.</td>
<td>Comparison Of Cognitive-Behaviour Therapy With Different Types Of Controls: Within-Group Mean Effect Size At Follow-Up</td>
</tr>
<tr>
<td>5.</td>
<td>Comparison Of Cognitive-Behaviour Therapy With Different Types Of Controls: Between-Group Mean Effect Size At Follow-Up</td>
</tr>
<tr>
<td>6.</td>
<td>Effect Sizes (At Follow-Up) Controlling For Treatment Dropout</td>
</tr>
<tr>
<td>7.</td>
<td>Re-calculated Between-Group Effect Sizes (At Follow-Up) Including Different Estimated Effects for Non-significant Results</td>
</tr>
<tr>
<td>8.</td>
<td>Effect Size (At Follow-Up) And Methodological Quality</td>
</tr>
<tr>
<td>9.</td>
<td>Effect Sizes (At Follow-Up) Controlling For Random Assignment</td>
</tr>
<tr>
<td>10.</td>
<td>Effect Sizes (At Follow-Up) Controlling For Integrity Control Of Treatment</td>
</tr>
</tbody>
</table>
List Of Figures

Figure 1. Between-Group Mean Effect Sizes For Comparison Of Cognitive-Behaviour Therapy With Different Types of Controls Across Follow-Up Periods. 52
Introduction

Cognitive-Behavioural Therapy (CBT) is a form of psychotherapy that combines cognitive-based and behaviour-based techniques in an effort to effect behaviour change (Beck, 1970; Ellis, 1962; Meichenbaum, 1977). Unfortunately, this simple definition falls short of being useful in practical application. Many practitioners and theorists have had difficulty throughout the evolution of CBT in defining the scope of CBT using this simple definition (Dobson & Block, 1988). The difficulty is not surprising since, according to some sources, CBT blurs the distinction between behaviour theory and cognitive theory (Grossberg, 1981; Phillips, 1981). To distinguish CBT from other forms of therapy, the theoretical origins of the techniques must be understood.

Proponents of classical behaviour therapy (BT) believe that behavioural theory explains and already incorporates into treatment the basic tenets of cognitively-based theories (Phillips, 1981). The behaviourists' claim that cognition has always been an integral part of BT has been expressed as follows:

...cognition is unavoidable in behaviour therapy, as it is in any form of psychotherapy and indeed, in almost all human activities....If behaviour therapists have not made a point of such things, it is because to do so is as redundant as to mention, when recording that a
patient received an intravenous injection, that a syringe was used. (Wolpe, 1978, p. 442-443) as cited in Grossberg 1981, p. 27.

These cognitive behaviours, however, are not believed to be functionally different from other behaviours, and thus, do not require different treatment (Phillips, 1981). As a consequence, proponents of BT describe cognitive-based theories and therapies as "retrogressive, misleading and even anti-scientific." (see Sweet, 1991, p. 159).

On the other hand, advocates of cognitive-based theories believe that their focus on the modification of maladaptive thoughts to bring about both cognitive and behaviour change using verbally-based therapies distinguishes cognitively-based theories, including CBT, from the mainly nonverbal means of BT (Miller & Berman, 1983). It is this focus on the alteration of faulty cognitions through cognitive restructuring (CR) that distinguishes CBT as a psychological treatment (Rachman & Wilson, 1980).

In a comparative review examining the degree of beneficial outcomes associated with BT versus CBT, Sweet (1991) chose to define CT and BT as separate entities, and combine the two definitions in an effort to define CBT. Referring to CT, Sweet (1991), states that:

...the essential identifying factor was that putative cognitions, cognitive processes, core beliefs, self statements, attitudes, attributions, schema, etc., were
being therapeutically addressed in a verbal fashion. The central assumptions of cognitive therapy is that the patients' wilful modification of these phenomena (via therapist instruction and assistance) yields significant changes in behavioural and physiological dependent variables as well as the cognitive variable themselves. (p. 161),

and continued by defining BT as:

...those procedures regularly and frequently called behavioural in the literature...systematic desensitization... applied relaxation, behavioural rehearsal, stimulus control, guided and unguided exposure in vivo, flooding...social skills training and assertiveness training, actively scheduling, self-monitoring, behavioural homework practice, response cost, contingent positive reinforcement, participant modelling, taken economies, time-out, covert conditioning and self-management strategies...

(p. 161)

Finally, the definition of CBT was simply stated by Sweet (1991), as CT combined with any one or more components of BT.

In contrast, Miller (1983) states that defining CBT as a combination of cognitive and behavioural techniques is theoretically sound but practically inappropriate. He maintains that many "CBT" treatments found in the research literature fail to report their behavioural components specifically. Hence, Miller (1983) proposed to define and identify CBT treatments as any form of treatment that includes the examination of
maladaptive beliefs. In discussing Miller's (1983) meta-analysis, Dush (1983), subsumed all cognitive-behavioural techniques reported in the overview under the label "cognitive restructuring". Dush (1983), distinguishes self-statement modification (SSM) from CR, and includes SSM in CBT's quiver of treatment techniques. In a meta-analysis reviewing the efficacy of self-statement modification, Dush (1983) describes Ellis' rational-emotional therapy, Meichenbaum's self-instructional therapy and Beck's CT as forms of CBT, "each emphasiz(ing) the importance of covert self-verbalizations and suggests that 'self-talk' or private monologues can influence performance of a wide variety of tasks." (p. 409). This 'self-talk' refers to adaptive or maladaptive self statements which are believed to be modificable.

The common theoretical component apparent in each of the operational definitions of CBT is the modification of maladaptive thoughts. This focus on maladaptive thoughts seems to be the component that distinguishes CBT from BT according to many if not all attempts at definition. In this meta-analysis, we have defined CBT as a therapy that involves the assessment of excesses in maladaptive thoughts or deficits in adaptive thoughts, and the modification of these self-statements by means of verbal consultation with a therapist, either through restructuring of the cognitions (challenging, disputation, replacement, correction of distortions) or through covert self-verbalization (rehearsal), combined with a BT treatment component such as systematic
Cognitive-Behaviour Therapy 5

desensitization, relaxation, rehearsal, stimulus control, exposure in vivo, flooding, social skills training, assertiveness training, scheduling, self-monitoring, behavioural homework practice, response cost, reinforcement, modelling, token economies, time-out, covert conditioning or self-management strategies. This definition and its focus on maladaptive thoughts reflects the core theoretical proposition of cognitive-based theories: that behaviour change is mediated by cognitive processes.

Cognitive-Behavioural Models And Strategies Of Problem Drinking

Cognitive-behavioural theories have been proposed to explain why people drink, where they drink, how much they drink and why most remain social drinkers while some fall victim to alcohol abuse and dependence (Wilson, 1987a). Although, many sources discuss the etiology of problem drinking in terms of genetic and biological factors, research evidence also supports the role of cognitively mediated variables in the development and maintenance of problem drinking (Wilson, 1987a).

The cognitive-behavioural models most often cited in the literature are the Tension-Reduction Theory and the Expectation Theory (Nathan, 1985; Oei, Lim, & Young, 1991; Wilson, 1987a; Wilson, 1987b). These theories are not unrelated. The basic tenet of the Tension Reduction Theory is that alcohol dependence is initially motivated by the need to reduce tension or stress (Oei, Lim, & Young, 1991; Wilson, 1987a). It is widely believed by problem drinkers themselves that alcohol reduces tension and
that by consuming alcohol the tension and stress will be decreased (Wilson, 1987b). However, studies over the past fifteen years suggest that alcohol consumption does not always result in emotional or physiological tension reduction (Oei, Lim, & Young, 1991; Wilson, 1987a; Wilson, 1987b). The pharmacological effects of alcohol have a variable impact on anxiety state (Wilson, 1987a).

The occasional tension-reducing effects of alcohol may produce reinforcement for drinking behaviour because the tension reduction occurs on an intermittent schedule of reinforcement (Oei, Lim, & Young, 1991). Hence, problem drinking results from the learned expectation (Expectation Theory) that stress and tension reduction may ensue, rather than from the actual physiological and affective effects of alcohol consumption. Peoples' expectations or beliefs regarding outcome are often better predictors of later behaviour than the actual consequences of their behaviour (Wilson, 1987b). Under the Expectation-Theory model, the problem drinker has learned a contingency relationship, expecting alcohol to reduce his or her psychological stress and physiological arousal (Oei, Lim, & Young, 1991). It is this tension-reduction expectation that is believed to be one of the major cognitive mediational components of alcohol consumption.

The cognitive-behavioural treatment strategies used by the studies examined in this review center upon cognitive-restructuring techniques (changing cognitive distortions)
combined with various behavioural procedures. The cognitive components of treatment focus on issues of self-efficacy, attributions (internal versus external) and outcome expectations (Wilson, 1987a; Wilson, 1987b). The cognitive procedures include techniques designed to increase patients' awareness of their own automatic thoughts (self-talk). They provide methods of altering distortions in an effort to replace them with more adaptive thoughts. The behavioural components are used to produce further therapeutic change (Wilson, 1987a; Wilson, 1987b). The performance procedures provide practice situations in which the problem drinker can gain experience and increased believability in his or her new, more adaptive cognitive responses (Greenberger & Padesky, 1995). Over time, the problem drinkers are believed to gain more confidence in these beliefs allowing them to replace previously held maladaptive cognitions, resulting in healthier behaviour.

Efficacy Of CBT In The Treatment Of Problem Drinking

Over the past two decades there has been a proliferation of literature on the efficacy of CBT in treatment generally (Miller & Berman, 1983; Dush, Hirt, & Schroeder, 1983; Sweet & Loizeaux, 1991; Phillips, 1981; Ledwidge, 1978; Kendall & Hollon, 1979) and in the treatment of problem drinking (Emrick, 1975; Emrick, 1974; Costello, Biever, & Baillargeon, 1977; Bien, Miller, & Tonigan, 1993; Saunders, 1989). These narrative and quantitative reviews do not, however, include an appreciable examination of the efficacy of CBT in the treatment of problem drinking, but focus
mainly on the treatment of depression and anxiety (Miller & Berman, 1983). The reviews on problem drinking examine the efficacy of other forms of therapy (Agosti, 1995), or comparisons of brief versus extended treatment (Bien, Miller, & Tonigan, 1993). This lack of attention to problem drinking in CBT efficacy reviews is not surprising, because primary studies evaluating the effectiveness of CBT in the treatment of problem drinking have only entered the literature over the past 15 years. Recently, however, a few reviews have been published that directly or indirectly explore the CBT and problem drinking research.

Oei, Lim and Young (1991) reviewed 13 empirical studies of CBT and substance abuse; 11 problem drinking studies and 2 methadone maintenance studies. CBT was found to be an effective therapeutic approach. Moreover, in reviewing studies which included CR as a treatment component, Oei et al. (1991) concluded that CR, in particular, was effective in the treatment of problem drinking. The review was originally designed as a meta-analysis, and criteria for including CBT papers were developed following criteria listed in Dush, Hirt and Schroeder (1983), and Miller and Berman (1983). A literature search produced 13 studies that fulfilled the requirements. The authors concluded, for no specified reason, that the combined sample was insufficient to conduct a meta-analysis. Therefore, their review was a qualitative review of 13 studies. Furthermore, Oei et al. (1991) included Social Skills Training (SST) and Stress Management
Training (SMT) under the scope of CBT treatments (defining SST and SMT as forms of treatment that "focus" on patient's maladaptive beliefs). Consequently, only 4 of the 11 problem drinking studies reviewed utilized CR as a component of therapy. Hence, the overall finding that CBT is effective in the treatment of problem drinking is based on 11 studies, 7 of which employ either SST or SMT as the only CBT treatment component, and the secondary finding that CR is particularly beneficial is based on a qualitative review of a small sample of four studies. However, as detailed below, many more relevant studies have since been conducted.

In a recent publication, Miller, Brown, Simpson, Handmaker, Bien, Luckie, Montgomery, Hester and Tonnigan presented a comprehensive review of outcome literature relating to alcohol treatment. The systematic search produced 211 studies which reported outcomes relating to alcohol treatment. Thirty different treatment modalities (each represented by 3 or more papers) were compared. No distinction is made by the authors between CT and CBT; they subsume both forms of the therapy under the treatment modality CT. Seven papers were grouped under this modality. Miller et al. (1995) found that "Cognitive Therapy" ranked tenth (out of 30) when compared to other therapies in the treatment of problem drinking, and conclude that this result is "encouraging" (p. 24).

Miller's extensive review is interesting and, at the very least, innovative. However, it is not a meta-analysis. The
basic methodology of the review follows a meta-analytic approach, but the final analysis or combining of outcomes does not observe meta-analytic procedures for statistical pooling principles. The drawback of the approach used by Miller et al. (1995) is that although the treatment modalities are rank ordered, no data are provided regarding the size of difference in efficacy among the treatment modalities. More importantly, the rankings are not based on the size of the treatment effect for each study but rather on whether or not an effect was present. The "pooled" result of the studies within a specific modality represents only a plus, minus or zero treatment effect combined with a methodology score. The reader is left with little bases for comparing treatment modalities, and no ability to assess the magnitude of effect of the specific treatment modalities.

Recently, two meta-analyses have been reported in the literature, by the same author, which evaluate the efficacy of various alcohol treatments in controlled trials (Agosti, 1994; Agosti, 1995). The meta-analyses differ in the type of outcome reported. In the first meta-analysis, Agosti (1994) found that only 3 of 15 studies reported significantly greater abstinence rates in the treated group compared to the control group at follow-up. In the second meta-analysis, comparing quantity consumption data in 12 controlled studies, Agosti's (1995) results revealed that treated individuals consumed significantly less alcohol than individuals in the control groups. The author suggests that abstinence may have been too conservative or
restrictive as an outcome measure as an explanation for this apparent disparity in the results of the two meta-analyses.

There are several significant limitations associated with the design and methodology of Agosti's (1994, 1995) meta-analyses. The most crucial of these is the fact that many studies which meet Agosti's inclusion criteria are not included in the analysis (; Miller, 1991; Skutle & Berg, 1987; Foy, Nunn, & Rychtarik, 1984; Alden, 1988; Sannibale, 1989). Moreover, the design, as Agosti (1995) acknowledges, fails to allow for differential analysis of separate treatment modalities. Finally, Agosti (1995) includes a study by Sanchez-Craig, Annis, Bronet, and MacDonald (1984) in which the same alcohol treatment program is provided to two groups differing only in controlled drinking versus abstinence treatment goals. This probably should not have been considered a controlled study. The limitations reduce the value of Agosti's meta-analyses.

Meta-Analysis

The most meaningful benefit of a rigorously performed meta-analysis of efficacy research is that it provides the reader with a quantitative estimate of the magnitude of effect of a certain therapy in the treatment of a specific malady. The analysis begins with the magnitude of treatment effect for each study and statistically pools the individual effects to produce an overall effect size (ES) for the pooled results. The ES's for specific treatment models can be compared.

Another benefit of a meta-analytic review is that the
results of individual studies with small samples can be pooled together in an effort to detect potentially important small effects (Miller & Berman, 1983). Also, study features that may differ across the individual studies (particularly methodological quality) may be statistically analyzed to determine their effect on the results (Miller & Berman, 1983).

Our aim in this meta-analysis is to determine the overall effectiveness of CBT in the treatment of problem drinking. We have included only controlled trials to help reduce the possibility of error from experimental bias. By focusing only on problem drinking, the review we provide should also avoid the possible confounding effects associated with inclusion of papers concentrating on addictions to other substances.

Method

Literature Search

Computerised, English-language literature searches were performed using Psyclit (Jan. 1974-Jun. 1995), Medline (Jan. 1966-Jun. 1995), Eric (Jan. 1966-Jun. 1995), Dissertation Abstracts (Jan. 1980-Dec. 1995), and Current Contents (Jul. 17/95-Aug. 14/95). We restricted these searches to studies based on human subjects in which the efficacy of Cognitive Behaviour Therapy in the reduction of alcohol intake of problem drinkers was studied. Our search strategies included MeSH terms such as Alcohol*; Alcoholism; Problem Drinking; Cognitive-behaviour; Cognitive Therapy; Cognitive Restructuring (see Appendix A for a full list of the MeSH terms employed). The literature suggests
that searches relying on MeSH terms can miss relevant papers due to the inconsistency of computerised database indexers (Wakeford & Roberts, 1993; Farbey, 1993). We therefore searched titles and abstracts in the computerised databases for relevant text words (identical or similar words to the MeSH terms found in Appendix A). Furthermore, names of prominent authors in the Alcohol-CBT field were used in searches on the Medline and Psyclit databases in an attempt to identify any remaining appropriate studies (see Appendix B for a full list of authors' names). Another strategy we employed to locate pertinent papers was perusal of the bibliographies of relevant overviews and reviews. Finally, in an effort to find unpublished studies (beyond searching the Dissertation Abstracts database), abstracts from relevant conferences were reviewed.

**Inclusion Criteria**

**Rationale**

In a review of the efficacy of various treatments for problem drinking, Saunders (1989) suggested the following criteria as important aspects of research to enable valid conclusions to be drawn: presence of comparison groups; random assignment; avoidance of extraneous treatments; adequate characterization of subjects; appropriate outcome variables; independent assessment of outcome; corroboration of self report; low attrition rate; and replicability of results (p. 123). Unfortunately, few studies can boast such strong methodological quality. Thus, authors of overviews must consider which
methodological variables are essential to obtaining valid results without inappropriately restricting the scope of the analysis. In examining a number of meta-analyses and semi-quantitative reviews which focused upon outcome studies regarding the efficacy of treatments (Agosti, 1995; Bien, Miller, & Tonigan, 1993; Dush, Hirt, & Schroeder, 1983; Miller & Berman, 1983; Oei, Lim, & Young, 1991; Sobell, Toneatto, & Sobell, 1990; Sweet & Loizeaux, 1991), the following criteria have emerged, and were adopted by us as critical elements for inclusion in the present meta-analysis:

1) At least one component of the therapeutic intervention must involve CBT as defined in the introduction (p. 6).

2) CBT must be compared with at least one other non-CBT group or control group (no-treatment or placebo-attention group).

3) The subjects must belong to the adult clinical population whose primary clinical complaint is directly associated with problem drinking.

4) The study must report data regarding the amount of alcohol consumption (amount consumed or days abstinent).

5) The subjects in the CBT group cannot be concomitantly using medication that may influence treatment outcome.

6) The study must be reported in the English language.

The photocopied papers were screened at this stage using these inclusion criteria. Two independent observers (Andrew Matthew (AGM) and Barry Schneider (BHS)) examined each paper, in its entirety, to determine if the inclusion criteria were met.
The agreement among the independent observers was assessed by the Kappa statistic (Fleiss, 1985). The observer agreement for acceptance/rejection was good (Kappa=0.88). Disagreement was resolved by consensus.

**Data Extraction**

All papers passing the inclusion criteria were reviewed (AM), and relevant information was extracted using a predesigned form. Data noted on this form included information regarding study specifics (e.g., year of publication), experimental characteristics (e.g., initial sample size), subject characteristics (e.g., age), therapist characteristics (e.g., experience level), cognitive behavioural treatment characteristics (e.g., total hours of therapy), and methodological quality (e.g., randomization) (see Appendix C for a full list of descriptive characteristics collected). Additional information needed to calculate the effect sizes, beyond that found in the papers, was solicited from 8 of the authors, with only moderate response (Rosenberg & Brian, 1986).

**Conventions Adopted**

Duplicate publications of the same trial were dealt with in two ways. If subsequent reports included a longer intervention we selected the most recent report. Otherwise, we chose the earliest publication and the later reports were used to supplement data as required (Ito, Donovan, & Hall, 1988). Where studies reported more than one alcohol consumption outcome measure, we calculated individual effect sizes for each drinking
measure and a single weighted mean average of these effect sizes was used in the analysis. In one study, (Ito, Donovan, & Hall, 1988), the sample sizes differed across measures so the average sample size was used for effect size weighting. Furthermore, some studies (Monti, Abrams, Binkoff, Zwick, & et-al, 1990; Brandsma, Maultsby Jr., & Welsh, 1980; Oei & Jackson, 1982) compared CBT treatment to more than one formal alternative treatment. The individual effect sizes for each alternative treatment condition were averaged to produce a single effect size per study for each treatment-control comparison. When alcohol consumption measures were reported that by definition cannot have baseline values (e.g., number of days to first drink, or number of days to first heavy drinking occasion), we included these outcomes in the analysis only if the study reported that the comparison groups did not differ at baseline in measured drinking intake (Monti, Abrams, Binkoff, Zwick, & et-al, 1990; Ito, Donovan, & Hall, 1988; Brandsma, Maultsby Jr., & Welsh, 1980).

Many studies reported follow-up results without reporting post-treatment results. One obvious reason for this is that many patients received the treatments in an inpatient setting without access to alcohol. The alcohol consumption outcomes upon immediate release to the community would be meaningless (Monti, Abrams, Binkoff, Zwick, & et-al, 1990; Oei & Jackson, 1984; Oei & Jackson, 1982). Nevertheless, in an effort to utilize the available information to its fullest, both post-treatment effect sizes (for those papers that reported post-treatment outcomes)
and follow-up effect sizes were calculated in this review. The follow-up effect sizes were calculated from studies reporting a minimum of 5.5 months follow-up. This period of follow-up was chosen for two reasons: shorter follow-up periods are likely to misrepresent relapse rates; a minimum of 5.5 months was the longest follow-up period most frequently reported in the papers reviewed. If more than one follow-up period was reported beyond 5 1/2 months (e.g., at 9, 12 and 24 months) then a single average effect size was calculated for the overall follow-up period (Kivlahan, Marlatt, Fromme, Coppel, & et-al, 1990; Baer, 1992; Brandsma, Maultsby Jr., & Welsh, 1980; Oei & Jackson, 1982). However, effect size was considered as a function of follow-up in a secondary analysis.

**Methodological Quality**

The methodological quality of each of the included studies was determined using Miller's et al. (1995) Methodological Quality Ratings Scale (see Appendix D). For each study, a single methodological quality score was calculated.

**Statistical Analysis**

In this meta-analysis, CBT was statistically compared to three separate types of control group. The *no-treatment group* consists of participants who do not receive any systematic treatment (e.g., waiting list, or assessment only). The *placebo group* participants receive only non-formal active treatment (e.g., unstructured discussion, attention only). Finally, the *alternative treatment group* members receive a formal active
treatment unrelated to CBT (e.g., relaxation therapy, social skills training).

The effects of CBT in the treatment of problem drinking were examined through both within-group and between-group effect size comparisons. The primary statistic used to evaluate treatment outcomes across trials and across groups was Hedges's $g$ (Hedges & Olkin, 1985), an unbiased effect size (E.S.) estimate. A Fixed Effect Size model was used because the focus of this meta-analysis was to evaluate whether or not CBT treatment has been effective, on the average, to date (Fleiss, 1993). The effect size, Hedges's $g$, was calculated using Schwarzer's meta-analysis software program version 5 (Schwarzer, 1989). The computerised program begins by calculating a standardized mean difference as proposed by (Glass, 1976):

$$E.S. = \frac{(X_0 - X_B)}{Pooled\ SD} - \frac{(X_0 - X_0)}{Pooled\ SD}$$

Pooled SD

where $X_0$ is the dependent measure outcome mean, and $X_B$ is the dependent measure baseline mean, and SD is the standard deviation. Therefore, E.S. represents a difference score of the change within each group; a quantitative description of the comparison of the effects (effect size) of each intervention (in some cases no intervention) over time. The pooled standard deviation used in this meta-analysis refers to the pooling of the baseline standard deviations for each treatment and control group. This differs from the more common methods of using the
control group standard deviation or the pooled within-group standard deviation (Glass, McGaw, & Smith, 1981; Smith, Glass, & Miller, 1980). The pooled baseline standard deviation standardizes the effect sizes in the same manner as using the control group standard deviation, and also helps to control for treatment and control group population variability in a similar manner as using a pooled within-group standard deviation. In this review, the advantage of using the pooled baseline standard deviation versus the pooled within-group standard deviation is that the statistics reported in the included studies were not always sufficient for the calculation of the latter. Thus, by using the pooled baseline standard deviations, fewer studies were excluded for statistical reporting reasons. The pooled baseline standard deviation was used to calculate both within-group and between-group effect sizes.

The resulting effect size for each comparison was weighted by sample size using the "weighted integration model" outlined in Schwarzer's meta-analysis program (Schwarzer, 1989. Version 5.0). Furthermore, Hedges and Olkin (1985, p. 80) show that this weighted effect size estimate has a small sample bias; therefore, to calculate the unbiased effect size (Hedges's 'g') the following formula was used:

\[ g = \left( 1 - \left( \frac{3}{4} \times N - 9 \right) \right) \times \text{weighted ES} \]

Most of the effect sizes were calculated from means and standard deviations. If these statistics were not reported in the study an attempt was made to contact the author(s) to obtain
missing information. Otherwise, the effect size was estimated from significance values. If a study reported the results as statistically significant without reporting the p-value, a p-value of 0.05 was assumed. Similarly, for reports of non-significant effects without a stated p-value, the p-value was assumed to be 0.5 (Mullen & Rosenthal, 1985).

In two studies, (Oei & Jackson, 1984; Rohsenow, Smith, & Johnson, 1985), a single significant within-group F-value was reported, combining the effects measured at several follow-up points. The p-value associated with the reported F-statistic was used to calculate the effect size for each follow-up period.

Directions were reversed where necessary so that a positive effect size indicated a reduction in alcohol consumption.

Results

Literature Search

The search strategies resulted in 219 citations. After the initial screening by one of the authors (AGM), 113 research papers remained (see appendix E for a list of papers excluded from the analysis at the initial screening stage). These 113 studies were then reviewed by independent observers (AGM and BHS). Of the 113 studies, 101 did not pass the inclusion criteria: 49 studies did not include a CBT treatment group, 9 reported on participants whose primary relevant problem was not associated with problem drinking, 23 did not have a control group, 9 did not report alcohol consumption outcome data, 4 were
review papers, 5 were not primary papers, and 2 were a duplicate publication of previously reported data (see Appendix F for a list of papers excluded from the analysis at the inclusion criteria stage). This left 12 studies for analysis (see references with asterisks in the reference list).

**Study Characteristics**

Information on the characteristics of the individual studies is provided in Table 1. A total of 480 subjects participated in the studies reviewed. The average number of hours of CBT treatment across the 12 studies was 15.3 hours with a median of 12 hours. All but one of the studies reported a follow-up period of at least 5.5 months (O'Malley et al., 1992). The average length of follow-up across studies was 9.8 months with a range of 3 months to 18 months. The majority of the included studies were randomized controlled trials. One study, (Ito, Donovan, & Hall, 1988), reported that while group assignment was not truly random, the subjects did not choose their own treatment, nor was any subject characteristic a determining factor for assignment. Rosenberg and Brian (1986) cited ethical concerns as a reason for failure to randomize (Rosenberg & Brian, 1986).

Four studies recruited their subjects from a alcohol treatment program (Ito, Donovan, & Hall, 1988; Monti, Abrams, Binkoff, Zwick, & et-al, 1990; Oei & Jackson, 1982; Oei & Jackson, 1984), 3 studies recruited from college campuses (Baer, 1992; Kivlahan, Marlatt, Fromme, Coppel, & et-al, 1990; Rohsenow, Smith, & Johnson, 1985), 2 studies recruited from the community
and outpatient clinics (Sannibale, 1989; O'Malley et al., 1992), and finally, 3 studies had subjects referred to them either by the courts or by medical doctors (Brandsma, Maultsby Jr., & Welsh, 1980; Robertson & et-al, 1986; Rosenberg & Brian, 1986).

Measures of alcohol consumption varied greatly across studies. Alcohol intake was reported in the number of Standard Ethanol Consumption (SEC) units consumed (Baer, 1992), the number of standard drinks consumed (Rosenberg & Brian, 1986; Kivlahan, Marlatt, Fromme, Coppel, & et-al, 1990), the number of ounces of ethanol consumed (Ito, Donovan, & Hall, 1988; Monti, Abrams, Binkoff, Zwick, & et-al, 1990; Oei & Jackson, 1982; Oei & Jackson, 1984; O'Malley et al., 1992; Brandsma, Maultsby Jr., & Welsh, 1980), and the number of grams of ethanol consumed (Robertson & et-al, 1986; Sannibale, 1989). Many of the studies differed in how this information was documented. Some of the studies reported consumption rates during the trial period (O'Malley et al., 1992), others reported intake at follow-up over 1 week (Oei & Jackson, 1982; Oei & Jackson, 1984), 1 month (Baer, 1992; Kivlahan, Marlatt, Fromme, Coppel, & et-al, 1990; Robertson & et-al, 1986; Sannibale, 1989; Rohsenow, Smith, & Johnson, 1985), 3 months (Brandsma, Maultsby Jr., & Welsh, 1980), and still others reported them during possible and actual drinking days (distinguished by the number of days a subject had access to alcohol, i.e., not in jail or detox) (Monti, Abrams, Binkoff, Zwick, & et-al, 1990). One study reported alcohol intake in amount consumed per hour per drinking occasion (Rosenberg &
Brian, 1986). Other methods of reporting alcohol consumption included: number of days drinking or abstinent (Monti, Abrams, Binkoff, Zwick, & et-al, 1990; O'Malley et al., 1992; Robertson & et-al, 1986; Brandsma, Maultsby Jr., & Welsh, 1980); number of binges (Brandsma, Maultsby Jr., & Welsh, 1980); Peak Blood Alcohol Level (PBAL) (Baer, 1992; Kivlahan, Marlatt, Fromme, Coppel, et-al, 1990); number of times heavy drinking (reported as 5 or more drinks (Rohsenow, Smith, & Johnson, 1985), or 6 or more drinks (Monti, Abrams, Binkoff, Zwick, & et-al, 1990) consumed during a single sitting); number of times of continuous drinking (reported as 12 or more hours of drinking) (Sannibale, 1989); number of drinking occasions in which the subject stopped intake at 2 or less drinks (a measure of controlled drinking) (Brandsma, Maultsby Jr., & Welsh, 1980); length of time to first drink (Ito, Donovan, & Hall, 1988; Monti, Abrams, Binkoff, Zwick, & et-al, 1990); and length of time to first heavy drinking episode (Monti, Abrams, Binkoff, Zwick, & et-al, 1990).

Although dissertation abstracts and unpublished papers were collected, only published journal articles met the inclusion criteria and are reviewed here. Nine of the 12 papers (Ito, Donovan, & Hall, 1988; Kivlahan, Marlatt, Fromme, Coppel, et-al, 1990; Monti, Abrams, Binkoff, Zwick, & et-al, 1990; Oei & Jackson, 1982; Oei & Jackson, 1984; O'Malley et al., 1992; Rohsenow, Smith, & Johnson, 1985; Rosenberg & Brian, 1986; Brandsma, Maultsby Jr., & Welsh, 1980) reported an integrity control measure of the treatments being delivered, while 7 papers
reported corroboration of subject self-report data (Ito, Donovan, & Hall, 1988; Oei & Jackson, 1982; Oei & Jackson, 1984; O'Malley et al., 1992; Robertson & et-al, 1986; Sannibale, 1989; Brandsma, Maultsby Jr., & Welsh, 1980) usually through the subject's significant other. Nine studies reported the experience level of the therapists. Seven of these 9 studies report employing at least one doctoral level therapist (Ito, Donovan, & Hall, 1988; Baer, 1992; Kivlahan, Marlatt, Fromme, Coppel, & et-al, 1990; O'Malley et al., 1992; Robertson & et-al, 1986; Rohsenow, Smith, & Johnson, 1985; Brandsma, Maultsby Jr., & Welsh, 1980).

Information on the subject characteristics for each of the studies is provided in Table 2. The overall weighted mean age of the sample across the 12 studies was 33 years, with the majority of the studies focusing on males. For the 7 studies that reported level of education achieved by the subjects, the mean is equal to 12 years. The authors endeavoured to extract data from the studies relating to the chronicity of the subjects' alcohol problems and the number of prior treatments, but these characteristics were rarely reported.

Overall Effect Sizes

Table 3 presents Hedges's (1985) unbiased effect sizes for each of the 12 studies. No effect size is indicated where the data were not available in the original source or additional data supplied by the authors. Because of the pattern of availability, the analysis from this point forward will focus on follow-up effect sizes. This should also be more relevant to clinical
practice in light of the common problem of relapse in the treatment of problem drinking. Available post-treatment effect sizes are also recorded in Table 3, though they were not analyzed statistically.

**Outlier Effect**

One study reported a post-treatment between-group effect size that was nearly eight times the magnitude of the mean weighted average of the other effect sizes reported (O'Malley et al., 1992). Moreover, this study's effect size is in the opposite direction from all other post-treatment between-group effect sizes. For these reasons, the between-group post-treatment effect size calculated from O'Malley's (1992) study (E.S. = -2.3025) can be treated as an outlier. Because, as mentioned above, we analyzed only the follow-up results, this outlier has not affected the statistical calculations of effect size herein.

**Overall Effects**

Table 4 presents the within-group effect sizes of within-group comparisons of CBT and the control groups. Following Cohen's (1977) guidelines, these results suggest that both CBT and alternative treatments were moderately effective in reducing the amount of alcohol consumption in problem drinkers. In contrast, the placebo controls appear to produce only a small effect, while the single no-treatment control produced an effect size close to zero. Mann-Whitney U statistics were calculated to compare the within-group effect sizes for CBT, alternative
treatments, and controls. There were no significant findings probably due to the small cell sizes.

Between-group analyses were also performed in order to summarize the direct comparisons within each study of the efficacy of CBT treatments and control treatments across all studies. In comparing CBT treatment with alternative treatments, 36 effect sizes were calculated and combined to produce 8 separate effect sizes, one per study. These 8 unbiased effect sizes were then pooled to produce an overall between-group weighted mean effect size of E.S. = .03 (see Table 5). Similarly, 18 effect sizes were combined to produce 6 individual study effect sizes which were statistically pooled for the comparison of CBT and placebo treatment. In the two studies examined to compare CBT treatment with no-treatment, 9 effect sizes were calculated across the different measures reported. These 9 effect sizes were then combined to produce the CBT versus no-treatment effect size reported in Table 5.

The near zero effect size for comparisons between CBT and alternative treatments suggests that CBT treatment does not differ in effectiveness in reducing alcohol consumption when compared to alternative treatments. There were small and moderate effect sizes, respectively, for CBT comparisons with placebo and no-treatment controls.

**Treatment Dropout**

For a problem drinking treatment program to be clinically or practically successful, it must not only produce a reduction in
participant alcohol consumption, but it must also avoid excessive participant dropout. In an effort to detect whether treatment dropout appreciably affected the effect sizes reported in this review, a specific analysis of treatment dropout was performed (see Table 6). Effect sizes were re-calculated assuming no effect for study subjects reported as treatment dropouts. It is clear from Table 6 that the total sample effect sizes do not appreciably differ from the effect sizes for treatment completers; therefore, treatment dropout had little influence on our results.

Different Methods For Calculating Effect Size From Non-Significant Results

It is quite common for studies to report some results as being "non-significant" but fail to report an exact p-value. The meta-analyst can choose to assign an effect size estimate of 0.00 to these "non-significant" studies; however, this conservative method is believed to lead to effect sizes that are too small (Rosenthal, 1995). Another method is to assign a p-value of .50 to these studies, as has been done in this meta-analysis. However, we believe that with such a small number of included studies in this meta-analysis, assigning a p-value of .50 is a good compromise between assigning an effect size of 0.00 and not including the "non-significant" variables. To make certain that this decision did not unduly affect our results, we re-calculated effect sizes using various methods of managing "non-significant" results. As detailed in Table 7, the results of the analysis
suggest that "non-significant" results had little impact on the magnitude of the effect sizes. The effect sizes corresponding to a p-value of .50 are not sizably different than effect sizes corresponding to other methods of administering "non-significant" results; hence, our decision to use a p-value of .50 had little impact, positive or negative, on the overall effect sizes.

Methodological Quality

Table 8 lists the effect sizes and methodological quality scores for each study. These scores were calculated using Miller's (1995) Methodological Quality Rating Scales. For both within-group and between-group comparisons, higher positive effect sizes are generally associated with higher ratings in methodological quality. The lowest quality score calculated (MQS = 7) (Robertson, 1986) was associated with the lowest within-group effect size, the lowest placebo treatment between-group effect size, and the third lowest alternative treatment effect size. Moreover, the 3 studies receiving the highest quality score calculated (MQS = 14) all reported positive effect sizes. A Spearman Coefficient was calculated to measure the correlation between effect size and methodological quality score. The resulting coefficient was positive \( r = .38 \) but non-significant.

Random Assignment

In order to determine what effects random assignment might have on effect size, a comparison was made between studies that reported randomly assigned subjects and those that did not (see Table 9). There were higher average effect sizes when subjects
were not assigned randomly to conditions. These increases in effect size were especially apparent in the alternative treatment pooled between-group comparison (delta ES = .64). Mann-Whitney U statistics were calculated for each comparison, but no significant differences were found.

Integrity Control of Treatment Delivered

Studies that reported some form of integrity control for the treatments delivered (e.g., video analysis, therapist-treatment manual) were compared with studies that did not report integrity control procedures (see Table 10). The pooled effect sizes associated with studies that included integrity controls were higher across both within-group and between-group comparisons. The greatest difference was reported for the comparison of CBT and alternative treatments (delta = .68). Mann-Whitney U statistics were calculated for each comparison, but no significant differences were found.

Follow-Up Interval

Figure 1 displays the relationship between weighted between-group effect sizes and length of follow-up. The no-treatment condition is not included in the figure due to too few data points. At post-treatment, the CBT treatment is moderately more successful in reducing participant alcohol consumption than the placebo treatment, however, this moderate effect is reduced to a small effect over time. In contrast, the small between-group effect size appears stable from post-treatment through to 6 months follow-up and then the effect size drops to near zero at
12 months. It should be noted that the 6 month results in the figure differ from the results reported in Table 5 because the effect sizes in the table are average effect sizes for periods of 5.5 months and greater.

Discussion

The results of both the within-group and between-group analyses suggest that CBT treatment is moderately more successful than no-treatment, slightly more successful than placebo treatment, and no more successful than other formal alternative treatments. These results are encouraging given the history of inconclusive results from treatment efficacy research in problem drinking (Saunders, 1989). Two reviews of problem drinking treatment outcome literature concluded not only that treatments did not differ in effectiveness, but that treatments possessed no benefit (Goodwin, 1988), or only slight benefit (Lindstrom, 1992) over no treatment. Our results, in contrast, are more promising, and are similar to those reported in more general reviews of the efficacy of CBT treatment (Miller & Berman, 1983; Dush, Hirt, & Schroeder, 1983), and in reviews that specifically investigate CBT treatment of problem drinking (Oei, Lim, & Young, 1991; Miller et al., 1995). Moreover, the consistency of our results across the within-group and between-group comparisons provides greater confidence in our outcomes.

The conclusions of this meta-analysis should be viewed conservatively because of the relatively small number of studies that qualified for inclusion in the examination. After reviewing
the inconclusive and controversial outcome literature associated with the treatment of problem drinking, we decided that this meta-analysis would therefore follow a "best evidence synthesis" approach as described by Slavin (Slavin, 1986; Slavin, 1987). Slavin (1986) proposes that the "best" available evidence in a field can be collected when systematic inclusion criteria are employed that focus on the substantive aspects and the methodological adequacy of individual studies. Furthermore, Slavin (1986) suggests that more information can be obtained from these higher quality studies than from the "statistical analyses of the entire methodologically and substantively diverse literature." (p. 7). The rigorous inclusion criteria set out in our methods section were designed to include studies with the highest internal and external validity. We believe the included studies in this meta-analysis represent the best evidence available in the literature to date.

We have paid specific attention to the definition of CBT treatment in an effort to clearly define the parameters of the treatment being evaluated. As referred to in the introduction, investigators and clinicians alike express difficulty in defining the distinctive qualities of CBT. While the reader may not agree with the definition of CBT set out in this review, nevertheless the definition provides a clear basis for the critical examination of our results and conclusions.

We also paid considerable attention to inclusion criteria for outcome measures. There are two schools of thought regarding
the goals of treatment for problem drinking. On the one hand, there are those that view abstinence as the only practical outcome, while others accept "controlled" or reductions in alcohol intake as a genuine goal of treatment (Saunders, 1989). We decided not to use abstinence as a criterion for the success of therapy because in a previous review, Agosti (1994) determined that it was unduly conservative for use in summarizing outcomes. Furthermore, abstinence data was not commonly reported. Alcohol consumption was the next obvious outcome measure to consider. Agosti (1995) also turned to alcohol consumption as the outcome measure for his most recent meta-analysis. In contrast to Agosti, however, we included abstinence measures (e.g., days abstinent over a specified period of time) as one of several measures of consumption which were combined to produce a single effect size for each study. We did not consider the effects of CBT on other dependent variables such as self-concept, locus of control or social condition (e.g., socioeconomic status, marital satisfaction). As in previous reviews (Agosti, 1995; Bien, 1993), we choose to focus on consumption as the pivotal outcome measure, based on our belief that, any other benefits of intervention are likely to be ephemeral unless consumption is reduced for an extended period of time.

Examination of treatment dropout was one of several secondary analyses we conducted. It is generally believed that subjects who fail to complete treatment and are not available for follow-up have likely returned to problem drinking (Nathan,
Failure to account for treatment dropout may therefore artificially exaggerate positive treatment effect. Our analysis failed to detect any appreciable influence on effect size when dropouts were included and zero treatment effect is assumed for them. There were very small dropout rates in the included studies in this secondary analysis, which probably accounts for the minimal effect of dropout. It should be noted that dropout rate was not reported in all studies; thus, it is possible that the true influence of dropout on the overall effect sizes may have been more substantial.

Similarly, an analysis was conducted to detect the differential impact of various methods of incorporating "non-significant" outcomes in this meta-analysis. We decided, following Mullen (1985), to use a p-value of .50 when calculating effect sizes for "non-significant" outcomes. As detailed in the results section, a more conservative approach is to assume an effect size of 0.00, however, because of the relatively small number of studies being analyzed we felt that this approach could obscure real outcomes. Our secondary analysis revealed that regardless of the method employed our results were not meaningfully influenced. The most likely reason for this lack of effect is that only one study reported "non-significant" results in the between-group comparison with alternative treatments (Brandsma, Maultsby Jr., & Welsh, 1980), and this was true for only two studies in comparisons with the placebo group (Baer, 1992; Rohsenow, Smith, & Johnson, 1985).
The average methodological quality score of the included studies in this meta-analysis was 11.3 out of a possible 17 points. Even with the "best evidence" approach, the included studies, on average, only rank in the upper middle portion of the Methodological Quality Scale. This suggests that the quality of research on the effectiveness of CBT in the treatment of problem drinking to date is only mediocre. The results of our analysis of methodological quality suggest that this finding may indeed have an impact on our conclusions. The Spearman Coefficient reported in the results section (p. 29), although non-significant, suggests the tendency toward a positive relationship between magnitude of effect size and study quality. These findings support the need for further, more methodologically sound research in the area. The reader should further note that this level of study quality is not restricted to CBT efficacy research, but is in fact quite common across various treatment modalities in the alcohol treatment outcome literature (Miller et al., 1995).

The results of our analysis of the impact of random assignment on effect size revealed a tendency (although non-significant) toward non-randomized studies producing larger effect sizes when compared to randomized studies. This suggests that the inclusion of non-randomized studies in our meta-analysis may have inflated our results. This effect was greatest for the between-group comparison with alternative treatments. These findings imply that future research in the quantitative synthesis
of literature in this area should control for the effects of random assignment.

Since definitions of CBT do vary, and, as noted above, often fail to provide clear direction to implementers of the intervention, another important issue was integrity control. Studies that reported some form of integrity control had a tendency (although non significant) to report higher effect sizes. The inclusion of studies that failed to report integrity control of treatment delivered may have had the impact of depressing our overall effect sizes. Given the definitional problems considered above, it is conceivable that the interventions delivered in the studies not featuring integrity controls might not be true applications of CBT. Once again, the impact on effect size was greatest for the between-group comparison with alternative treatments. Thus, some aspects of methodological quality were associated with higher effect sizes, while other aspects were associated with smaller effects sizes. These conflicting setting-by-treatment interactions are not uncommon (Pillemer & Light, 1980).

Our analysis of effect size as a function of length of follow-up revealed a small positive effect of CBT treatment over alternative treatment at 6 months follow-up. Only at 12 months follow-up did the relative equality of CBT and alternative treatment become evident. The importance of at least a 12 month follow-up when comparing treatment outcomes is suggested by these results.
The series of study design and reporting differences discussed above force the meta-analysts to make subjective decisions about inclusion/exclusion and methods of quantitative pooling which may influence the results of the meta-analysis. In an effort to reduce the number of these types of decisions, and to increase the systematic nature of the study collection pooling techniques, we recommend a number of specific study design features and reporting characteristics for future primary research:

1) Report means and standard deviations to ensure easy and efficient calculation of effect sizes estimates. (Also we recommend that researchers report means and standard deviations whenever possible, even for non-significant results. The large sample size of many quantitative summaries can detect trends in data that smaller individual studies are unable to detect).

2) Standardize the measure of alcohol intake across studies so that the results can be easily compared. The most common measure of alcohol intake found in the studies reviewed in this meta-analysis was ounces of alcohol consumed per day. This measure is simple and easily combined or compared.

3) Report the number of subjects that dropped out of each treatment during the treatment period, and if possible report "treatment completer" results and "total sample" results.
4) Randomly assign subjects to treatment and control groups where possible.

5) Give full descriptions of treatments being delivered in the study, and include a measure of the integrity of the delivery of those treatments.

6) Include a minimum 12 month follow-up in the study design. Results of this meta-analysis discussed earlier suggest that a 6 month follow-up is not long enough to detect true long-term outcomes.

Conclusion

In 1976, Aaron T. Beck stated that in order for cognitive therapy to be considered a form of psychotherapy it had to provide evidence through empirical research and treatment outcome studies of its effectiveness (Beck, 1976). Since then, there has been an abundance of outcome research conducted that supports the effectiveness of CT and CBT, especially in the treatment of depression and anxiety (Dobson, 1989; Miller & Berman, 1983; Sokol, Beck, Greenberg, Berchick, & Wright, 1989). Although the results of this meta-analysis are encouraging, it is essential for further high quality outcome research to be performed and analyzed before CBT can be as highly regarded in the treatment of problem drinking as it is in the treatment of depression and anxiety.
## Experimental Characteristics

<table>
<thead>
<tr>
<th>Author and Year (Country)</th>
<th>N</th>
<th>Treatment</th>
<th>Hours</th>
<th>Length of Follow-up (Months)</th>
<th>Random Assignment</th>
<th>Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Baer 1992 (USA)</td>
<td>67</td>
<td>Advice</td>
<td>9</td>
<td>12-24</td>
<td>Yes</td>
<td>14</td>
</tr>
<tr>
<td>2. Brandsma 1980 (USA)</td>
<td>63</td>
<td>Dynamic</td>
<td>30</td>
<td>12</td>
<td>Yes</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Alcoholics Anon</td>
<td></td>
<td>No Treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Ito 1988 (USA)</td>
<td>34</td>
<td>Dynamic</td>
<td>12</td>
<td>6</td>
<td>No</td>
<td>11</td>
</tr>
<tr>
<td>4. Kivlahan 1990 (USA)</td>
<td>36</td>
<td>Information</td>
<td>12</td>
<td>12</td>
<td>Yes</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No Treatment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Monti 1990 (USA)</td>
<td>53</td>
<td>Skills Training</td>
<td>12</td>
<td>6</td>
<td>Yes</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Skills Training</td>
<td>with Family</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td>Oei 1982</td>
<td>19</td>
<td>Behaviour Therapy</td>
<td>24</td>
<td>12</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>(New Zealand)</td>
<td></td>
<td>Discussion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td>Oei 1984</td>
<td>16</td>
<td>Behaviour Therapy</td>
<td>24</td>
<td>6</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>(New Zealand)</td>
<td></td>
<td>with Discussion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td>O'Malley 1992</td>
<td>31</td>
<td>Supportive Therapy</td>
<td>12</td>
<td>3</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>(USA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9.</td>
<td>Robertson 1986</td>
<td>33</td>
<td>Advice</td>
<td>12</td>
<td>15.5</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>(Scotland)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.</td>
<td>Rohsenow 1985</td>
<td>28</td>
<td>Advice</td>
<td>6</td>
<td>5.5</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>(USA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11.</td>
<td>Rosenberg 1986</td>
<td>20</td>
<td>Behaviour Therapy</td>
<td>24</td>
<td>6</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>(USA)</td>
<td></td>
<td>Advice</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12.</td>
<td>Sannibale 1989</td>
<td>80</td>
<td>Confrontation</td>
<td>7</td>
<td>15.5</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>(Australia)</td>
<td></td>
<td>Advice</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note.** *No methodological quality score was calculated for O'Malley (1992) because no follow-up results were reported; thus, the follow-up criteria in the Methodological Quality Scale were not applicable. CBT = Cognitive-Behaviour Therapy.*

*Scores calculated using Miller's (1995) Methodological Quality Scale (see Appendix D).*
Table 2

Subject Characteristics

<table>
<thead>
<tr>
<th>Author and Year Published</th>
<th>Description of Subjects</th>
<th>Age Mean/Min/Max</th>
<th>Sex (Male)</th>
<th>Employment</th>
<th>Mean Education (Years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Baer 1992</td>
<td>MAST-1 problem</td>
<td>21/17/40</td>
<td>48%</td>
<td>Students</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>BAL &gt;.10% twice/week</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Brandsma 1980</td>
<td>Alcoholic</td>
<td>46/24/58</td>
<td>100%</td>
<td>69%</td>
<td>10</td>
</tr>
<tr>
<td>3. Ito 1988</td>
<td>Inpatient Treatment</td>
<td>37/-/-</td>
<td>100%</td>
<td>36%</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>ADS = 20</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Kivlahan 1990</td>
<td>ADS = Maximum of 13</td>
<td>23/18/35</td>
<td>58%</td>
<td>Students</td>
<td>12</td>
</tr>
<tr>
<td>5. Monti 1990</td>
<td>DSM-III Dependence</td>
<td>43/24/71</td>
<td>100%</td>
<td>12%</td>
<td>11</td>
</tr>
<tr>
<td>6. Oei 1982</td>
<td>Inpatient Treatment</td>
<td>32/-/-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7. Oei 1984</td>
<td>Inpatient Treatment</td>
<td>32/-/-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>8. O'Malley 1992</td>
<td>DSM-III-R Dependence</td>
<td>38/-/-</td>
<td>-</td>
<td>73%</td>
<td>-</td>
</tr>
<tr>
<td>9. Robertson 1986</td>
<td>SADQ = 14.1</td>
<td>36/-/-</td>
<td>80%</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Shortened MAST = 5.6
<table>
<thead>
<tr>
<th>Study</th>
<th>Group Type</th>
<th>BAL</th>
<th>Abstinent</th>
<th>Partial</th>
<th>Complete</th>
<th>Age Range</th>
<th>% Male</th>
<th>% Female</th>
<th>% Other</th>
<th>Other</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>10. Rohsenow 1985</td>
<td>Heavy Social Drinkers</td>
<td>21/20/24</td>
<td>100%</td>
<td>Students</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Rosenberg 1986</td>
<td>DUI Offenders</td>
<td>30/-/-</td>
<td>100%</td>
<td>87%</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Sannibale 1989</td>
<td>ADS = 11.2, MAST = 17</td>
<td>32/-/-</td>
<td>100%</td>
<td>58%</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note.** Dashes in the table represent missing data that was not available from the original sources. MAST = Michigan Alcohol Screening Test; BAL = Blood Alcohol Level; ADS = Alcohol Dependence Scale; SADQ = Severity of Dependence Score; DUI = Driving Under the Influence.
Table 3

Effect Sizes for Individual Studies

<table>
<thead>
<tr>
<th>Author and Year</th>
<th>Published</th>
<th>Within-Group ES (for CBT)</th>
<th>Between-Group ES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Post-treatment (n)</td>
<td>Follow-up (n)</td>
</tr>
<tr>
<td>1. Baer 1992</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2. Brandsma 1980</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>1.04 (15)</td>
<td>-</td>
</tr>
<tr>
<td>3. Ito 1988</td>
<td>-</td>
<td>.62 (14)</td>
<td>-</td>
</tr>
<tr>
<td>4. Kivlahan 1990</td>
<td>.54 (15)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5. Monti 1990</td>
<td>-</td>
<td>.63 (16)</td>
<td>-</td>
</tr>
<tr>
<td>6. Oei 1982</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7. Oei 1984</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>8. O'Malley 1992</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Study</td>
<td>Effect Size</td>
<td>Sample Size</td>
<td>Outcome</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>-------------</td>
<td>-------------</td>
<td>---------</td>
</tr>
<tr>
<td>9. Robertson 1986</td>
<td>1.27</td>
<td>(14)</td>
<td>-</td>
</tr>
<tr>
<td>10. Rohsenow 1985</td>
<td>.26</td>
<td>(11)</td>
<td>PL .23</td>
</tr>
<tr>
<td>11. Rosenberg 1986</td>
<td>.05</td>
<td>(7)</td>
<td>ALT .17</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PL .97</td>
</tr>
<tr>
<td>12. Sannibale 1989</td>
<td>.09</td>
<td>(41)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-</td>
</tr>
</tbody>
</table>

Note. ALT = Alternative Treatment; PL = Placebo Treatment; NTX = No Treatment.
Table 4
Comparisons of Cognitive-Behaviour Therapy With Different Types of Controls: Within-Group Mean Effect Size at Follow-Up

<table>
<thead>
<tr>
<th>Treatment</th>
<th>n of Group Studies</th>
<th>Mean Within-Group ES</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBT</td>
<td>7</td>
<td>.50</td>
<td>.05 to 1.27</td>
</tr>
<tr>
<td>Alternative</td>
<td>5</td>
<td>.69</td>
<td>.27 to .95</td>
</tr>
<tr>
<td>Placebo</td>
<td>4</td>
<td>.33</td>
<td>.14 to .61</td>
</tr>
<tr>
<td>No Treatment</td>
<td>1</td>
<td>.04</td>
<td></td>
</tr>
</tbody>
</table>
Table 5

Comparison of Cognitive-Behaviour Therapy With Different Types of Controls: Between-Group Mean Effect Size at Follow-Up

<table>
<thead>
<tr>
<th>Type of Control</th>
<th>n of Studies</th>
<th>Mean Between-Group ES (SE)</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alternative</td>
<td>8</td>
<td>.03 (.12)</td>
<td>-.47 to 1.68</td>
</tr>
<tr>
<td>Placebo</td>
<td>6</td>
<td>.22 (.13)</td>
<td>-.60 to 1.23</td>
</tr>
<tr>
<td>No Treatment</td>
<td>2</td>
<td>.44 (.27)</td>
<td>.36 to .59</td>
</tr>
</tbody>
</table>
Table 6

Effect Sizes (at Follow-Up) Controlling for Treatment Dropout

<table>
<thead>
<tr>
<th>Within-Group (for CBT)</th>
<th>n of Studies</th>
<th>ES</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment Completers&lt;sup&gt;a&lt;/sup&gt;</td>
<td>5</td>
<td>.59</td>
<td>.05 to 1.04</td>
</tr>
<tr>
<td>Total Sample</td>
<td>5</td>
<td>.51</td>
<td>.05 to .87</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Between Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alternative Treatment</td>
</tr>
<tr>
<td>Treatment Completers&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Total Sample</td>
</tr>
<tr>
<td>Placebo Treatment</td>
</tr>
<tr>
<td>Treatment Completers&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Total Sample</td>
</tr>
</tbody>
</table>

<sup>a</sup>Based only on studies where dropout rate was reported, in order to facilitate comparison.
Table 7

Re-calculated Between-Group Effect Sizes (at Follow-Up) Including Different Estimated Effects for Non-Significant Results

<table>
<thead>
<tr>
<th>Basis of Estimate</th>
<th>Alternative Treatment</th>
<th>Placebo Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n of Studies</td>
<td>ES (SE)</td>
</tr>
<tr>
<td>ES = 0</td>
<td>8</td>
<td>-.02 (.12)</td>
</tr>
<tr>
<td>Not included</td>
<td>7</td>
<td>-.02 (.14)</td>
</tr>
<tr>
<td>p = .50</td>
<td>8</td>
<td>.02 (.12)</td>
</tr>
<tr>
<td>p = .10</td>
<td>8</td>
<td>.09 (.12)</td>
</tr>
</tbody>
</table>

Note. "Not included" refers to the set of effect sizes for the respective studies when "non-significant" variables were left out of the calculations.
<table>
<thead>
<tr>
<th>Author and Year Published</th>
<th>MQS</th>
<th>Within-Group ES (n)</th>
<th>Between-Group ES (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Baer 1992</td>
<td>14</td>
<td>-</td>
<td>PL .16 (67)</td>
</tr>
<tr>
<td>2. Brandsma 1980</td>
<td>14</td>
<td>-</td>
<td>ALT .23 (50)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NTX .35 (35)</td>
</tr>
<tr>
<td>3. Ito 1988</td>
<td>11</td>
<td>1.04 (15)</td>
<td>ALT .18 (30)</td>
</tr>
<tr>
<td>4. Kivlahan 1990</td>
<td>10</td>
<td>.62 (14)</td>
<td>ALT .31 (26)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NTX .57 (24)</td>
</tr>
<tr>
<td>5. Monti 1990</td>
<td>11</td>
<td>.63 (16)</td>
<td>ALT -.29 (53)</td>
</tr>
<tr>
<td>6. Oei 1982</td>
<td>12</td>
<td>-</td>
<td>ALT 1.05 (14)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PL 1.05 (14)</td>
</tr>
<tr>
<td>7. Oei 1984</td>
<td>11</td>
<td>-</td>
<td>ALT 1.50 (16)</td>
</tr>
<tr>
<td>8. O'Malley 1992</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>8. Robertson 1986</td>
<td>14</td>
<td>1.27 (14)</td>
<td>PL .83 (33)</td>
</tr>
<tr>
<td>Study Description</td>
<td>Sample Size</td>
<td>Effect Size</td>
<td>Notes</td>
</tr>
<tr>
<td>-------------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------</td>
</tr>
<tr>
<td>Rosenberg 1986</td>
<td>7</td>
<td>.05 (7)</td>
<td>ALT -.22 (13)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PL -.56 (13)</td>
</tr>
<tr>
<td>Sannibale 1989</td>
<td>11</td>
<td>.09 (41)</td>
<td>ALT -.46 (72)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PL -.06 (59)</td>
</tr>
</tbody>
</table>

**Note.** No Methodological Quality Score was calculated for O'Malley (1992) because the study only reported post-treatment results; therefore, the follow-up criteria in the Methodological Quality Ratings Scale was not applicable. MQS = Methodological Quality Score; ALT = Alternative Treatment; PL = Placebo Treatment; NTX = No Treatment.
Table 9

**Effect Sizes (at Follow-Up) Controlling for Random Assignment**

<table>
<thead>
<tr>
<th></th>
<th><strong>n of Studies</strong></th>
<th><strong>ES</strong></th>
<th><strong>Range</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Within-Group (for CBT)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random Assignment</td>
<td>5</td>
<td>.45</td>
<td>.09 to 1.27</td>
</tr>
<tr>
<td>No Random Assignment</td>
<td>2</td>
<td>.71</td>
<td>.05 to 1.04</td>
</tr>
<tr>
<td><strong>Between-Group</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Alternative Treatment</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random Assignment</td>
<td>4</td>
<td>-.14</td>
<td>-.47 to .32</td>
</tr>
<tr>
<td>No Random Assignment</td>
<td>4</td>
<td>.50</td>
<td>-.24 to 1.68</td>
</tr>
<tr>
<td><strong>Placebo Treatment</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random Assignment</td>
<td>4</td>
<td>.21</td>
<td>-.06 to .99</td>
</tr>
<tr>
<td>No Random Assignment</td>
<td>2</td>
<td>.25</td>
<td>-.60 to 1.23</td>
</tr>
</tbody>
</table>
Table 10

Effect Sizes (at Follow-Up) Controlling for Integrity Control of Treatment

<table>
<thead>
<tr>
<th></th>
<th>n of Studies</th>
<th>ES</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within-Group (for CBT)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Integrity Control</td>
<td>5</td>
<td>.60</td>
<td>.05 to 1.04</td>
</tr>
<tr>
<td>No Integrity Control</td>
<td>2</td>
<td>.39</td>
<td>.09 to 1.27</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>n of Studies</th>
<th>ES</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between-Group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alternative Treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Integrity Control</td>
<td>7</td>
<td>.21</td>
<td>-.29 to 1.68</td>
</tr>
<tr>
<td>No Integrity Control</td>
<td>1</td>
<td>-.47</td>
<td></td>
</tr>
<tr>
<td>Placebo Treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Integrity Control</td>
<td>3</td>
<td>.25</td>
<td>-.60 to 1.23</td>
</tr>
<tr>
<td>No Integrity Control</td>
<td>4</td>
<td>.21</td>
<td>-.06 to .99</td>
</tr>
</tbody>
</table>
Figure 1. Between-group weighted mean effect sizes for comparisons of cognitive-behaviour therapy treatment with both alternative and placebo treatments, across posttreatment and two separate follow-up periods. ALT = Alternate Treatment; PL = Placebo Treatment.
References

*Indicates studies included in the meta-analysis.


MeSH Terms Employed In The Computerised Search Strategies


Cognitive-beh* (C-), Cognitive The* (theory and therapy), C- Techniques, C- Restructuring, C- Mediation, C- Affective-Behavioural, C- Based, C- Behavioural-Action, C- Behavioural Affective, C- Behavioural Psychodynamic, C- Behavioural Systems, C-Behavioural Therapy, C- Clinical, C- Coping, C- Distortion, C- Emotional-Behavioural, C- Emotive Behavioural, C- focused, C- Imagery, C- Intervention, C- Oriented, C-Psychology, C-Social Learning, C-Social Behavioural, CBT.

Rational-Emotive Therapy, Rational-Behaviour Therapy, RET.

Appendix B

Authors Individually Searched In Medline and Psyclit

Annis, H.M.                  Botvin, G.J.
Brown, H.P.                  Connors, G.J.
Donovan, D.M.                Eriksen, L.
Foy, D.W.                    Heather, N.
Miller, W.R.                 Monti, P.M.
Oei, T.P.                    Rohsenow, D.J.
Rosenberg, H.                Sanchez-Craig
Sobell, L.C.                 Sobell, M.B.
Werch, C.E.                  Wilson, G.T.
Appendix C

Descriptive Data Extracted From Individual Studies.

Study Specifics:
Authors
Year of publication
Country study performed
Journal

Experimental Characteristics:
Type of report [journal, dissertation, (unpublished)]
Treatments evaluated (cognitive only, cognitive and non-cognitive)
Controls used (placebo, waiting list, assessment only)
Length of run-in period
Length of wash-out period
Length of follow-up
Sample size at baseline
Sample size at treatment outcome (attrition)
Sample size at follow-up (attrition)
Alcohol intake at baseline (mean, standard deviation and/or error)
Alcohol intake at treatment outcome (abstinence, quantity)
Alcohol intake at follow-up
Number of study centres

Subject Characteristics:
Age (min, max, mean)
Sex (number of males, number of females)
Appendix C continued

Race (percentage non-white)
Education level
Employment
Source of subjects (advertisements, college campus, community etc.)
Diagnosis of subjects (DSM-III/R/IV abuse, DSM-III/R/IV dependence, other)
Inpatient or outpatient population
Number of previous treatments
Chronicity of subject (years of problem drinking)

Therapist Characteristics:
Sex (male, female, both)
Experience level (some graduate experience, Doctoral student, doctoral, other)
Profession (Psychology, Psychiatry, Social Work)

Cognitive-Behavioural Treatment Characteristics:
Principle orientation (Beck, Ellis, Meichenbaum)
Number of sessions
Total hours of therapy
Span of treatment (weeks)
Treatment modality (group, individual, combination)
Cognitive treatment components (restructuring, self-statement modification)

Behavioural treatment components (systematic desensitization, etc.)
Goal of treatment (abstinence versus controlled)

Methodological Quality:

Randomization
Quality control of treatments delivered
Personal versus questionnaire follow-up
Type of corroboration (urinalysis, informant confirmation)
Accounting for treatment attrition
Accounting for follow-up attrition
Blinding of outcome assessment interviewing
Appendix D

**Methodological Quality Scale** (Miller et al. 1995)

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Group Allocation</td>
<td>4</td>
<td>Randomization</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Within-subjects counterbalanced design</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Case control, matching, alternative cohorts</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>Quasi-experimental design</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>Violated randomization or nonequivalent groups</td>
</tr>
<tr>
<td>B. Quality Control</td>
<td>1</td>
<td>Treatment standardized by manual, procedures, specific training</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>No standardization specified</td>
</tr>
<tr>
<td>C. Follow-up Rate</td>
<td>2</td>
<td>85-100% follow-ups completed</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>70-84.9% follow-ups completed</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>&lt; 70% follow-ups completed</td>
</tr>
<tr>
<td>D. Follow-up Length</td>
<td>2</td>
<td>12 months or longer</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>6-11 months</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>&lt; 6 months</td>
</tr>
<tr>
<td>E. Contact</td>
<td>1</td>
<td>Personal or telephone contact for &gt; 70% of completed follow-ups</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>Questionnaire, unspecified, or &lt; 70%</td>
</tr>
<tr>
<td>F. Collaterals</td>
<td>1</td>
<td>Collaterals interviewed</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>No collateral verification</td>
</tr>
</tbody>
</table>
### Cognitive-Behaviour Therapy 67

**Appendix D continued**

<table>
<thead>
<tr>
<th>G. Objective</th>
<th>1 = Objective verification (records, serum, breath)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 = No objective verification</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>H. Dropouts</th>
<th>1 = Treatment dropouts are enumerated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 = Dropouts neither discussed nor accounted for</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>I. Attrition</th>
<th>1 = Cases lost to follow-up are enumerated and considered in outcome reporting</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 = Lost cases not enumerated or considered in outcome reporting</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>J. Independent</th>
<th>1 = Follow-up done by treatment-blind interviewer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 = Follow-up non-blind; not specified; questionnaire</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>K. Analyses</th>
<th>1 = Appropriate statistical analyses of group differences are reported</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 = No statistical analyses; clearly inappropriate analyses</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>L. Multi-site</th>
<th>1 = Parallel replications at two or more sites, with separate research teams</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 = Single site or comparisons of sites offering different programs</td>
</tr>
</tbody>
</table>
Appendix E

The references found in this appendix did not meet the initial screening criteria and were excluded from the meta-analysis. The references are listed under bolded headings citing specific reasons for exclusion.

No CBT Treatment Group


Appendix E continued


Appendix E continued


Appendix E continued


Appendix E continued


The Primary Clinically Relevant Problem Was Not Problem Drinking


Appendix E continued


**Editorial/Commentary Papers**

Appendix E continued


Case Reports


Appendix E continued


**Papers Not Reported In English**


Appendix F
The references found in this appendix did not meet the inclusion criteria and were excluded from the meta-analysis. The references are listed under bolded headings citing specific reasons for exclusion.

No Cognitive-Behaviour Therapy Group Involved In The Trial


Appendix F continued


Appendix F continued


Appendix F continued


Appendix F continued


The Primary Clinically Relevant Problem Was Not Problem Drinking


Appendix F continued


Appendix F continued


Appendix F continued


No Alcohol Consumption Outcome Measures


Appendix F continued


**Review Papers (no original data collection)**


**Not Primary Reports**


Appendix F continued

Duplicate Publications
