Clinical Effectiveness of a Cognitive Behavioral Group Treatment Program for
Anxiety Disorders: A Benchmarking Study.

Tian P.S. Oei^{a,b} and Mark J. Boschen^{c,*}

^aSchool of Psychology, The University of Queensland, Brisbane, Australia,

oei@psy.uq.edu.au

^bCBT Unit, Toowong Private Hospital, Brisbane, Australia.

^cSchool of Psychology, Griffith University, Gold Coast, Australia,

m.boschen@griffith.edu.au, Ph: +61 7 55528283, Fax: +61 7 55528291.

*Corresponding Author

Abstract

Previous research has established efficacy of cognitive behavioral therapy (CBT) for anxiety disorders, yet it has not been widely assessed in routine community clinic practices. Efficacy research sacrifices external validity to achieve maximum internal validity. Recently, effectiveness research has been advocated as more ecologically valid for assessing routine clinical work in community clinics. Furthermore, there is a lack of effectiveness research in group CBT. This study aims to extend existing research on the effectiveness of CBT from individual therapy into group therapy delivery. It aimed also to examine outcome using not only symptom measures, but also measures of related symptoms, cognitions, and life quality and satisfaction. Results from a cohort of patients with various anxiety disorders demonstrated that treatment was effective in reducing anxiety symptoms to an extent comparable with other effectiveness studies. Despite this, only 43 percent of individuals showed reliable change, and 17 percent were 'recovered' from their anxiety symptoms, and the post-treatment measures were still significantly different from the level of anxiety symptoms observed in the general population.

KEYWORDS: Anxiety, Anxiety Disorders, Benchmarking, Cognitive Therapy, Cognitive Behavior Therapy

Clinical Effectiveness of a Group Treatment Program for Anxiety Disorders:

A benchmarking study.

1.0 Introduction

Anxiety disorders are chronic, high-prevalence conditions (Kessler, Berglund, Demler, Jin, & Walters, 2005) that result in significant impairment of functioning and quality of life (Olatunji, Cisler, & Tolin, 2007). Disorders of anxiety have attracted increasing interest in the research literature over the last few decades, including large numbers of treatment trials (Boschen, 2008a, 2008b). Controlled trials and meta-analyses offer strong support for the efficacy of cognitive behavioral interventions for anxiety disorders such as specific phobia (Norton & Price, 2007), panic disorder (Mitte, 2005; Oei, Llamas, & Devilly, 1999), obsessive compulsive disorder (Eddy, Dutra, Bradley, & Westen, 2004), post-traumatic stress disorder (Benish, Imel, & Wampold, in press; Bisson & Andrew, 2007), social phobia (Taylor, 1996), and generalized anxiety disorder (Covin, Ouimet, Seeds, & Dozois, 2008). Such treatments compare favorably when contrasted with other treatment options such as psychopharmacological agents (e.g., Bandelow, Seidler-Brandler, Becker, Wedekind, & Rüther, 2007; Eddy et al., 2004).

1.1 The Need to Establish Effectiveness of Treatments

Despite clear evidence for efficacy of these interventions, several authors have questioned how readily such research findings generalize to routine clinical practice (e.g., Borkovec & Castonguay, 1998; Goldfried & Wolfe, 1998; Persons, 1991). Studies conducted for the purpose of establishing efficacy of a psychological treatment have typically focused on achieving maximum internal validity by employing a number of strict exclusion criteria (Borkovec & Castonguay, 1998;

Chambless & Hollon, 1998). One common exclusion criterion is the presence of comorbid disorders. Exclusion of individuals with significant psychopathology in addition to the primary treatment target is aimed at obtaining a relatively ‘pure’ measure of the ability of the intervention to exert a treatment effect on the disorder under scrutiny. Presence of comorbid conditions, however, is particularly common in the anxiety disorders (Kessler, Chiu, Demler, & Walters, 2005). Furthermore, such comorbidity is known to result in poorer prognosis and response to treatment (e.g., van Balkom, van Boeijin, Boeke, van Oppen, Kempe, & van Dyck, 2008). As such, the results of efficacy studies may exclude a large sub-population for whom treatments may be less effective, reducing the generalizability of their results.

Another common exclusion criterion in carefully controlled efficacy studies is adjunctive use by participants of psychopharmacological treatments. Adjunctive pharmacotherapy is viewed in efficacy studies as a potential confounding variable, potentially adding statistical noise and obscuring psychotherapy effects. Although this allows for well-controlled research, it is not in keeping with evidence that the overwhelming majority of anxiety disorder patients are prescribed medication for their condition. Some research estimates that between 55 and 90 percent of anxiety disorder patients utilize pharmacological agents to assist in treatment of their condition (Wardle, 1990).

Other concerns center on the potential differences between how treatment is conducted in research settings versus routine clinical practice. Numerous authors cite a number of factors that may differ between clinical practice and controlled treatment trials (e.g., Kendall & Southam-Gerow, 1995; Westbrook & Kirk, 2005). These include differences in adherence to standard manualized treatment protocols,

differences in level of supervision, and differences in levels of specialist treatment skills applicable to the treated condition.

With these questions in mind, there is increasing awareness of the need to demonstrate that the findings of efficacy studies are similar when treatments are utilized in routine clinical care environments (Borkovec & Castonguay, 1998; Chambless & Hollon, 1998; Goldfried & Wolfe, 1998; Westbrook & Kirk, 2005). Evidence has emerged over the past two decades that when treatments are conducted in routine clinical practice settings, effect sizes are more modest than those observed in controlled research environments (e.g., Weisz, Donenberg, Han, & Weiss, 1995). More recently, Westbrook and Kirk (2005) have asserted that effectiveness research characterizes a new era of more ecologically valid psychotherapy research.

In the largest clinical effectiveness study conducted to date, Westbrook and Kirk (2005) examined treatment outcome of a group of over 1200 individuals with a variety of anxiety, mood, and eating disorders treated in the British National Health Service. The authors noted a host of limitations to their dataset, consistent with the clinical rather than research setting. Despite these limitations, the authors observed large effects of treatment, particularly when floor-effects were taken into account. Furthermore, approximately half of the patients who completed treatment demonstrated reliable and clinically significant change. These results were compared with research-based benchmarks, with the authors concluding that although substantial, their treatment effect sizes were somewhat lower than those observed in previous research trials. Similar outcomes have been observed in other studies with smaller samples (e.g., McEvoy & Nathan, 2007; Persons, Bostrom, & Bertagnolli, 1999; Stuart, Treat, & Wade, 2000; Westbrook & Hill, 1998). There is preliminary

evidence also that treatment effects from routine clinical practice persist for at least two years post-treatment (e.g., Stuart et al., 2000; Westbrook & Hill, 1998).

Use of group-based interventions has been established in controlled trials for many anxiety (e.g., Braga, Cordoli, Niederauer, & Manfro, 2005, Erickson, Janeck, & Tallman, 2007) and mood disorders (e.g., Oei & Dingle, 2008). Group-based interventions potentially offer several advantages over individual treatments such as increased cost-effectiveness (Tucker & Oei, 2007) and increased access to treatments (Oei & Dingle, 2008). As for individual therapy, there is a need to establish whether the outcomes seen in research are similar to those seen in routine clinical group-based practice of cognitive behavior therapy.

A recent study using group CBT for the treatment of social anxiety disorder in a sample of 153 individuals, provided support for generalizability of treatments beyond the research environment (McEvoy, 2007). In this study, the author reported that the observed effect sizes in his community clinic were similar to those seen in previous efficacy studies. Furthermore, over half of those individuals who completed treatment achieved reliable symptom change, with one third reaching clinically significant symptom reduction. Similar results have also been reported by Gaston, Abbott, Rapee, and Neary (2006) in another group social phobia treatment study, and Rosenberg and Hougaard (2005) in a sample of individuals with panic disorder who underwent group CBT.

1.2 The Importance of Convergent Measures

One limitation of previous effectiveness research has been the use of a limited range of outcome variables. The recent effectiveness study by Westbrook and Kirk (2005) reported on a large sample of individuals treated in clinical practice, but was limited in its assessment of outcome. In this study, only the Beck Anxiety Inventory

and the Beck Depression Inventory were used to evaluate the effectiveness of treatment. In the assessment of treatment of anxiety disorders, the Beck Anxiety Inventory provides a reliable and valid measure of the physiological symptoms of anxiety, yet it fails to adequately assess other domains known to be important in anxiety disorders. Constructs such as anxiety-related cognitions, and depression symptoms and cognitions are important in outcome. Measurement of cognitive processes and content during cognitive therapy provides evidence of the processes of change that relate to the treatment. Similarly, patient quality of life and life-satisfaction provide ecologically valid outcomes that are often neglected (McAlinden & Oei, 2006). When outcomes show convergence between reduced symptoms, reduced negative cognitions, and increased life quality and satisfaction, this provides compelling evidence of the positive outcome of treatment, beyond simple reductions in symptom intensity.

1.3 The Current Study

The current study aims to further previous research into the generalizability of research-based treatment efficacy into routine clinical practice. It describes the treatment of a sample of hospital outpatients treated using group CBT for anxiety disorders between 1995 and 2002. Although the dataset was gathered before the integration of newer cognitive techniques (e.g., mindfulness), this allows comparison with other previous effectiveness studies with datasets of a similar age (e.g., Westbrook & Kirk, 2005). It extends on previous work such as that of Westbrook and Kirk (2005) by looking at group treatment in a large sample, and on previous group treatment studies such as that by McEvoy (2007), by examining group treatment in a larger sample of diverse anxiety disorder presentations. Results from the group treatment program are benchmarked against previous anxiety disorder efficacy

studies, as well as effectiveness studies with individually treated anxiety disorder patients. Attempts were made to measure a collection of relevant variables in addition to anxiety symptoms, such as anxiety-cognitions, depression symptoms and cognitions, quality of life, and satisfaction with life.

2.0 Method

2.1 Participants

Patients completed a group CBT program between 1995 and 2002 in the CBT Unit at a Private Hospital in Brisbane, Australia. A total of 396 patients were referred to the program by their treating psychiatrists for treatment of an anxiety disorder. Specific diagnostic information was available for 179 of the participants from the cohort, with 30.2% being diagnosed with panic disorder, 14% being diagnosed with generalized anxiety disorder, 8.4% having a primary diagnosis of posttraumatic stress disorder, and 17.3% being referred with a primary major depressive disorder with clinically significant anxiety symptoms. In keeping with the demands of routine clinical care settings, no patient was excluded either on the basis of having a comorbid disorder or their medication status. A total of 260 of the patients were female, 88 were male, and no sex was recorded for 48 other patients in the group. Ages ranged from 15 to 72 with a mean age of 42.60 years ($SD = 12.32$). Information about level of formal education was recorded for 268 of the participant group: Of these, 13.4% had not completed high school, 31.3% had completed high school, 26.5% had completed a diploma or certificate course, 14.6% had completed an undergraduate degree, and 14.2% had completed a postgraduate degree.

2.2 Materials

A collection of psychometric instruments was utilized to assess a broad array of variables, including psychological symptoms and psychosocial functioning. The

primary outcome measure (see below for a description of the Beck Anxiety Inventory) was selected as a psychometrically valid assessment of anxiety symptoms that was applicable across a range of different anxiety disorders. Other measures were selected to assess specific fears, cognitions associated with anxiety, and overall life quality and satisfaction. Measures were administered during the first and final group therapy sessions, and were collected by the group therapists for later review. Therapists did not review the measures for missing data until a later date. Where up to two individual questions from any measure were missing, these were replaced with the mean of other items. There was no more than one outlier on any of the measures except for the Mobility Inventory Accompanied subscale (two outliers), and the Satisfaction with Life Scale (five outliers). No individual participant was a consistent outlier on any more than one scale. As such, outlier cases were not excluded from further analysis. Not all questionnaires were given to all groups, resulting in missing data on all measures. Cronbach's alpha reliability figures for each of the measures in our sample are presented in Table 1.

Beck Anxiety Inventory (BAI; Beck, Epstein, Brown, & Steer, 1988). The BAI is a 21-item measure of the severity of physiological anxiety symptoms over the previous week. The BAI has good internal consistency ($\alpha = .92$) and item-total correlations that range from .30 to .71. Test-retest reliability over a one week period is .75. The BAI also has been shown to have good concurrent and discriminant validity, being able to discriminate homogenous and heterogeneous anxious diagnostic groups from other psychiatric groups (Beck et al., 1988).

Catastrophic Cognitions Questionnaire – Modified (CCQ-M; Khawaja, Oei, & Baglioni, 1994). The CCQ-M is based on Beck's cognitive theory of anxiety stipulating that cognitive themes of threat and danger are an important maintaining

factor in anxiety disorders. The CCQ-M has 21 items that are rated according to how dangerous each item is believed to be. The questionnaire is validated for use in both clinical and non-clinical samples. The measure shows high internal consistency, with Cronbach's alpha scores above .85 for all three subscales. Test-retest reliability over a two-week interval is good.

Depression Anxiety Stress Scales (DASS; Lovibond & Lovibond, 1995). The DASS is a self-report inventory of depression, anxiety and stress symptoms. It is designed to maximally discriminate between these three constructs. The DASS is available in a 21-item short form, which was used in the current study. Only the Stress subscale of the 21-item DASS was used. The DASS has demonstrated reliability and validity for use in clinical populations.

Fear Questionnaire (FQ; Marks & Matthews, 1979). The FQ comprises three phobia subscales including agoraphobia, blood-injury, and social. The survey also includes a question asking respondents to list the main phobia they want treated and other situations to which they are phobic. The FQ has been demonstrated to be appropriate for use with clinical populations (Oei, Moylan, & Evans, 1991).

Mobility Inventory for Agoraphobia (MIA; Chambless, Caputo, Jasin, Gracely & Williams, 1985). The MIA was specifically designed to measure self-reported severity of agoraphobic avoidance behavior. The MIA consists of 27 items asking the degree to which the places or situations are avoided due to discomfort or anxiety. Subjects are asked to rate the degree they avoid each place or situation for when they are accompanied and when they are alone. The MIA has been found to be a stable and internally consistent scale that is sensitive to change following treatment (e.g., Kwon, Evans, & Oei, 1991).

Panic and Agoraphobia Scale (PAS; Bandelow, 1999). The PAS has 13 items designed to measure the severity of panic disorder and agoraphobia symptoms. The PAS has five subscales (panic attacks, phobic avoidance, anticipatory anxiety, disability, and worries about health) with two or three items in each. The scale shows good internal consistency and validity for use in anxiety disorder samples.

Quality of Life Inventory (QOLI; Frisch, Cornell, Villanueva, & Ratzlaff, 1992). The QOLI is a self-report measure designed for use in clinical populations to measure quality of life across a range of different domains. The total unweighted score of all 16 items is used throughout the study reported in this paper. The QOLI has been demonstrated to be reliable and valid for use in an anxiety disorder population (McAlinden & Oei, 2006).

Satisfaction With Life Scale (SWLS; Diener, Emmons, Larsen, & Griffin, 1985). The SWLS was designed to assess an individual's overall judgment of their life in order to measure the concept of life satisfaction. The scale consists of five items.

Zung Self Rating Scale for Depression (Zung SRSD; Zung, 1965). The Zung SRSD is a 20-item self-report measure of depressive symptoms, designed for use in clinical populations. It has established reliability and validity for clinical use.

2.3 Procedure

The group CBT treatment was delivered on site at a private hospital in Brisbane, Australia. Each group consisted of between 8 and 14 individual participants. Two four-hour treatment sessions per week were conducted, over a total of four weeks, yielding 32 hours of group contact. The manualized treatment protocol consisted of several modules, including psychoeducation, group discussions, role-play exercises, arousal management training, cognitive restructuring, and exposure

planning. Our treatment program was similar to Westbrook and Kirk (2005) in that it used patients treated before (1987-1998) the rapid expansion of new cognitive therapy techniques such as mindfulness. Within the group, treatment was individualized through tasks such as individual cognitive restructuring homework and assistance with development of individual exposure hierarchies. Homework exercises, and information sheets were also utilized. Each group was run by a team of two clinical psychologists and a psychiatric nurse, each of whom was experienced in running CBT groups. The clinical psychologists had postgraduate clinical psychology degrees, and at least one was a senior clinical psychology academic with over 30 years experience in delivery of cognitive behavioral interventions.

3.0 Results

3.1 Equivalence of Completers and Non-Completers

Completers were defined as those participants for whom pre-treatment and post-treatment BAI scores were available, while non-completers were those participants for whom pre-treatment BAI scores were available, but no post-treatment BAI measure was collected. Comparisons between completers and non-completers on demographics and pre-treatment symptom measures were conducted to ensure that the completers were representative of the overall sample. Comparisons between completers and non-completers are reported in Table 1, revealing no significant differences between completers and non-completers, apart from the completers beginning treatment with a significantly better self-reported life satisfaction.

3.2 Treatment Significance and Effect Size

Treatment effect sizes were estimated for each outcome measure using the formula $(M_{\text{Pre}} - M_{\text{Post}}) \div SD_{\text{Pre}}$. Similarly to Westbrook and Kirk (2005), we acknowledge that such an estimate assumes no change without therapy. We have

elected to use this to enable comparisons with Westbrook and Kirk (2005). Effect sizes and significance tests are presented in Table 2.¹ Uncontrolled effect sizes for unselected patients ranged from 0.12 for the Zung SRSD to 0.69 for the DASS Stress subscale. Changes on each outcome measure over the course of treatment were statistically significant, except for the Zung SRSD.

We examined the subset of our patient sample that started with the most severe anxiety symptoms as measures by the BAI. Using a similar approach to that of Westbrook and Kirk (2005), we examined those individuals with a pre-treatment BAI score of at least 20. Effect sizes were generally larger in this subsample, with a BAI treatment effect size of 1.17, and an effect of 0.95 on scores from the DASS Stress subscale. Results of analyses in this subsample are presented in Table 3.

3.3 Clinical Significance

To allow for comparisons with the results reported by Westbrook and Kirk (2005), we adopted the same criteria used in their study. For our individual patients to be evaluated as ‘recovered’ they needed to meet four criteria: First, to be included, data from pre-treatment and post-treatment BAI scores was needed. Second, they needed to commence treatment with a BAI score of 11 or greater. Third, (to establish reliable change) patients were required to show a drop of 10 points on the BAI. Fourth, (to establish clinically significant change) the patient’s final BAI score was required to be 10 or less. These criteria were based on those suggested by Jacobsen and colleagues (Jacobsen & Truax, 1991; Jacobsen, Roberts, Berns, & McGlinchey, 1999). Data was available for 159 patients who met the first two criteria for this analysis. Of these a total of 68 (42.8%) showed reliable change, while 27 (17.0%) showed clinically significant improvement. Two patients (1%) showed reliable deterioration.

3.4 Normative Comparisons

Normative comparisons assess whether the treated group's symptoms are reduced sufficiently that they are indistinguishable from the level seen in the normal population (Kendall et al., 1999). Normative comparisons for clinical significance follow a series of five steps as outlined by Kendall et al. (1999). We used the BAI as our primary outcome measure and compared this to the normative data reported by Gillis, Haaga, and Ford (1995). Only patients for whom a pre-treatment and post-treatment BAI score was available were used in this analysis. Two separate comparisons were made: firstly using the entire sample of participants for whom pre-treatment and post-treatment BAI scores were available ($n = 181$), and secondly for the subset of patients with severe anxiety, defined as pre-treatment BAI scores of at least 20 ($n = 62$).

For the overall sample, the pre-treatment mean was 25.60 ($SD = 12.81$), while the post-treatment BAI scores had a mean of 17.39 ($SD = 13.06$). The normative sample mean reported by Gillis et al. (1995) in a sample of 267 was 6.6 ($SD = 8.1$). For Step 1, we followed the convention of defining scores of up to one standard deviation above the normative mean as clinically equivalent to normal, yielding a δ_1 of -8.1. In Step 2, the difference between the means of the clinical group and the normative sample were compared against the lower limit of the range specified by the δ_1 value from Step 1. This comparison demonstrated that the treated group was not clinically equivalent to the normative sample ($_{CE}t[446] = 2.69, p < .01$). In Step 3 we conducted a traditional test of the difference between the mean of the post-treatment group and the normative sample. This showed that the treated group remained statistically different from the normal population ($_{trad}t[446] = 10.76, p < .01$). On the basis of these two statistical tests, in Step 4 we categorized the findings as falling in

Kendall et al.'s Cell III (i.e., that the treated group was not clinically equivalent, and statistically different from the normative sample). Step 5 of the procedure of Kendall et al. (1999) is the graphical depiction of results, as shown in Figure 1.

For the severely anxious group ($BAI \geq 20$), the pre-treatment mean BAI score was 35.24 ($SD = 10.28$) and the post-treatment mean BAI score was 32.18 ($SD = 10.28$). This subsample was neither clinically ($_{CET}[114] = 10.27, p < .01$) nor statistically ($_{trad}t[327] = 21.15, p < .01$) equivalent to the normative sample after treatment. The results of the clinical equivalency analyses are shown in Figure 1.

3.5 Benchmarking

To assess comparative effectiveness of the group CBT program, comparisons were made between our effect size and that reported by previous authors. For comparison we chose Westbrook and Kirk (2005) as the largest effectiveness study for treatment of anxiety disorders, as well as research by McEvoy and Nathan (2007) and Rosenberg and Hougard (2005) as the two other studies of effectiveness of group treatment for anxiety disorders. The lack of diagnostic information prevented direct comparisons with previous studies that have investigated the effectiveness of CBT in the treatment of specific diagnoses. Similarly to Westbrook and Kirk (2005), effect sizes were not as large as those reported by Chambless and Gillis (1993) in their review of efficacy studies. They were, however, comparable to those reported by Westbrook and Kirk. Comparisons with previous effectiveness studies using cognitive behavioral interventions for anxiety disorders are presented in Table 4.

4.0 Discussion

The current study extends previous research into outcomes of anxiety disorder treatment in several ways. First, it contributes to the increasing volume of research examining the effectiveness of CBT when used under routine clinical practice

conditions. Second, it extends existing research on the effectiveness of CBT from individual therapy into group therapy delivery. Third, it assesses effectiveness using not only symptom measures, but also measures of related symptoms, cognitions and life quality and satisfaction.

The primary finding to emerge was the reduction of anxiety symptoms over the course of treatment. Moderate to large effect sizes of treatment on the primary outcome measure (according to the criteria of Cohen, 1992) were observed in the overall sample and in a subsample of the most severe cases. Effect sizes were roughly equivalent to that reported by Westbrook and Kirk (2005) in their individual treatment. They did, however, fall short of the very large effect sizes reported in previous efficacy studies (e.g., Chambless & Gillis, 1993). Such findings suggest that therapy performed in routine practice, with clients suffering from comorbid conditions, may yield effects less impressive than when used with carefully selected client populations typical of efficacy studies.

Despite the evidence for statistically significant change in outcome variables, more careful examination of the results reveals some limitations of treatment. In our cohort only approximately 17 percent of individuals treated in the group program were recovered by the end of treatment. These results suggest that many individuals experience a reduction in anxiety symptoms that falls short of full recovery. Previous effectiveness studies such as that reported by Westbrook and Kirk (2005) have documented recovery rates of between 25 and 35 percent for patients undergoing individual treatment. Our results suggest that group treatment, while effective, may result in fewer patients reaching full recovery from their anxiety symptoms.

Another method of assessing clinical improvement was comparison of the treated cohorts to the level of anxiety symptoms observed in the normal population.

In the present study the group did not return to a level of anxiety symptoms that made them indistinguishable from the normal population. These results further support the notion that while treatment may result in statistically significant improvement in symptoms, many individuals treated for anxiety disorders experience incomplete remission of symptoms, with a residual level of anxiety that still distinguishes them from those without a history of anxiety disorder.

One important finding from this study is that not only was the group CBT program effective at reducing psychopathology and associated cognitive content, but also at improving life quality and satisfaction. Effect sizes of around 0.3 were observed on these two variables. These findings suggest that the impact of the treatment extended beyond reduced levels of anxiety symptoms, but also led to significant improvement in perceived well-being.

The use of a heterogeneous sample of individuals with a range of different anxiety conditions also has relevance to the growing body of literature on transdiagnostic treatment approaches. Recent treatment studies (e.g., Norton, 2008; Norton, Hayes, & Hope, 2004) have demonstrated the efficacy of a transdiagnostic approach to treatment of anxiety disorders in groups. Such treatment methods are also consistent with generalized models of formulation of anxiety disorders (e.g., Boschen & Oei, 2008). Our results support the evidence of Norton et al. for a treatment approach to anxiety which focuses on the common elements of the conditions.

4.1 Limitations and Future Directions

When interpreting the findings of the current study, it is important to be mindful of several limitations of the current research. Limitations of the current study fall broadly into three categories: limitations of the available dataset, limitations of the

assessment methods used to evaluate outcome, and limitations of the treatment approach used.

Our dataset was limited in several ways. One limitation was that of missing data. The nature of clinical practice means that the near-complete datasets often seen in efficacy studies are not available. Our dataset had considerable missing data, and it is unclear how a more complete dataset may have altered our findings. A second dataset limitation is that there were a number of patients who dropped out from the treatment program. We did not use methods such as carrying forward the last observation and using this as the outcome data. Furthermore, several potential confounding variables were not available for analysis. Information regarding specific diagnoses, medication status, and external life events would have enhanced the current work. Extending assessment to include other methods such as clinician rating, and behavioral assessment would also be beneficial for future researchers engaging in similar work.

The assessment methodology of the current study was also limited in some respects. We utilized the BAI as the primary outcome measure. While the BAI is a psychometrically valid measure of physiological arousal symptoms associated with anxiety, it may have failed to assess for change in more specific symptoms associated with different anxiety disorders. Gaston et al. (2006), for example, found a pre-post effect size of 1.0 in a group of individuals treated for social phobia when outcome was assessed with the more specific Social Interaction Anxiety Scale. It is not possible in the current study to assess whether change in more specific variables may have yielded larger effect sizes than those observed on the more general BAI measure. Another potential limitation of our assessment method is the different time frames specified in each of the questionnaires. While some (e.g., the BAI) specify that the

questions are to be answered on the basis of the person's experience over the past week, others (e.g., the FQ) do not specify a time period. This may introduce problems when different time periods are being rated. It should also be recognized that in a treatment program delivered over a four-week period, changes in some domains (e.g., quality of life) may lag behind changes in symptoms.

Finally, limitations in our treatment approach should be acknowledged. Although treatment adhered to a standard treatment protocol, there were no checks made of treatment fidelity, beyond those of normal clinical supervision. This is similar to the study by Westbrook and Kirk (2005), in which no data regarding the details of the treatment program were available. Also, with our data being gathered between 1995 and 2002, it could also be argued that the treatment used may not reflect current CBT practices, including the incorporation of newer techniques such as mindfulness skills. Despite this, our cohort was treated with a standard CBT approach that is similar to that which is still likely to be seen in routine practice in many areas, and allows comparison with previously published studies from the same period. The reader should, however, be cautious in comparing our results to more recent studies in which newer cognitive techniques were utilized.

There remains considerable room for research into effectiveness of cognitive behavioral therapy for anxiety disorders in the conditions seen in routine clinical practice. Future research may use improved methodology, such as distinguishing between different anxiety disorders, and using measures that are more specific to the anxiety disorders under investigation. With a relatively small number of patients showing recovery, and the cohort showing residual symptoms that continued to distinguish them from the non-anxiety disordered population, it is clear also that there remains considerable room for improvement in treatment of anxiety disorders.

4.2 Conclusions

The current study adds to the growing body of evidence that cognitive behavioral interventions for the anxiety disorders are not only efficacious, but also effective when conducted in routine clinical practice. Furthermore, cognitive and behavioral treatments for anxiety disorders are effective when administered in a group therapy format, with groups containing a diverse range of anxiety problems. In addition to reducing anxiety symptoms, treatment also leads to an improvement in other related variables such as anxiety-related cognition, satisfaction with life, and quality of life.

References

- Bandelow, B. (1999). *Panic and Agoraphobia Scale (PAS)*. Seattle: Hogrefe & Huber Publishers.
- Bandelow, B., Seidler-Brandler, U., Becker, A., Wedekind, D., & Rüther, E. (2007). Meta-analysis of randomized controlled comparisons of psychopharmacological and psychological treatments for anxiety disorders. *World Journal of Biological Psychiatry*, 8, 175-187.
- Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology*, 56, 893-897.
- Benish, S. G., Imel, Z. E., & Wampold, B. E. (in press). The relative efficacy of bona fide psychotherapies for treating post-traumatic stress disorder: A meta-analysis of direct comparisons. *Clinical Psychology Review*.
- Bisson, J., & Andrew, M. (2007). Psychological treatment of post-traumatic stress disorder (PTSD). *Cochrane Database of Systematic Reviews*, 18, CD003388.
- Borkovec, T. D., & Castonguay, L. G. (1998). What is the scientific meaning of empirically supported therapy? *Journal of Consulting and Clinical Psychology*, 66, 136-142.
- Boschen, M. J. (2008a). Publication trends in individual anxiety disorders: 1980-2015. *Journal of Anxiety Disorders*, 22, 570-575.
- Boschen, M. J. (2008b). The growth of PTSD in anxiety disorder research. *Psychiatry Research*, 158, 262-264.
- Boschen, M. J., & Oei, T. P. S. (2008). A cognitive behavioural case formulation framework for treatment planning in anxiety disorders. *Depression and Anxiety*, 25, 811-823.

- Braga, D. T., Cordioli, A. V., Niederauer, K., & Manfro, G. G. (2005). Cognitive-behavioral group therapy for obsessive-compulsive disorder: a 1-year follow-up. *Acta Psychiatrica Scandinavica*, 112, 180-186.
- Chambless, D. L., Caputo, G. C., Bright, P., & Gallagher, R. (1984). Assessment of 'fear of fear' in agoraphobics: The Body Sensations Questionnaire and the Agoraphobic Cognitions Questionnaire. *Journal of Consulting and Clinical Psychology*, 52, 1090-1097.
- Chambless, D. L., & Gillis, M. M. (1993). Cognitive therapy of anxiety disorders. *Journal of Consulting and Clinical Psychology*, 61, 248-260.
- Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology*, 66, 7-18.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155-159.
- Covin, R., Ouimet, A. J., Seeds, P. M., & Dozois, D. J. (2008). A meta-analysis of CBT for pathological worry among clients with GAD. *Journal of Anxiety Disorders*, 22, 108-116.
- Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction With Life Scale. *Journal of Personality Assessment*, 49, 71-75.
- Eddy, K. T., Dutra, L., Bradley, R., & Westen, D. (2004). A multidimensional meta-analysis of psychotherapy and pharmacotherapy for obsessive-compulsive disorder. *Clinical Psychology Review*, 24, 1011-1030.
- Erickson, D. H., Janeck, A. S., & Tallman, K. (2007). A cognitive-behavioural group for patients with various anxiety disorders. *Psychiatric Services*, 58, 1205-1211.
- Frisch, M. B., Cornell, J., Villanueva, M., & Retzlaff, P. J. (1992). Clinical validation of the Quality of Life Inventory: A measure of life satisfaction for use in

- treatment planning and outcome assessment. *Psychological Assessment*, 4, 92-101.
- Gaston, J. E., Abbott, M. J., Rapee, R. M., & Neary, S. A. (2006). Do empirically supported treatments generalize to private practice? A benchmark study of a cognitive-behavioural group treatment program for social phobia. *British Journal of Clinical Psychology*, 45, 33-48.
- Gillis, M. M., Haaga, D. A. F., & Ford, G. T. (1995). Normative values for the Beck Anxiety Inventory, Fear Questionnaire, Penn State Worry Questionnaire, and Social Phobia and Anxiety Inventory. *Psychological Assessment*, 7, 450-455.
- Goldfried, M. R., & Wolfe, B. E. (1998). Toward a more clinically valid approach to therapy research. *Journal of Consulting and Clinical Psychology*, 66, 143-150.
- Jacobsen, N. S., Roberts, L. J., Berns, S. B., & McGlinchey, J. B. (1999). Methods for defining and determining the clinical significance of treatments effects: Description, application, and alternatives. *Journal of Consulting and Clinical Psychology*, 67, 300-307.
- Jacobsen, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59, 12-19.
- Kendall, P. C., Marrs-Garcia, A., Nath, S. R., & Sheldrick, R. C. (1999). Normative comparisons for the evaluation of clinical significance. *Journal of Consulting and Clinical Psychology*, 67, 285-299.
- Kendall, P. C., & Southam-Gerow, M. A. (1995). Issues in the transportability of treatment: The case of anxiety disorders in youths. *Journal of Consulting and Clinical Psychology*, 63, 702-708.

- Kessler, R. C., Berglund, P., Demler, O., Jin, R., & Walters, E. E. (2005). Lifetime prevalence and the age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, 62, 593-602.
- Kessler, R. C., Chiu, W. T., Demler, O., & Walters, E. E. (2005). Prevalence, severity, and comorbidity in 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, 62, 617-627.
- Khawaja, N. G., Oei, T. P. S., & Baglioni, A. J. (1994). Modification of the Catastrophic Cognitions Questionnaire (CCQ-M) for normals and patients: Exploratory and LISREL analysis. *Journal of Psychopathology and Behavioral Assessment*, 16, 325-342.
- Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety and Stress Scales with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy*, 33, 335-343.
- Marks, I. M., & Matthews, A. M. (1979). Brief standard self-rating for phobic patients. *Behaviour Research and Therapy*, 17, 263-267.
- McAlinden, N. M., & Oei, T. P. S. (2006). Validation of the Quality of Life Inventory for patients with anxiety and depression. *Comprehensive Psychiatry*, 47, 307-314.
- McEvoy, P. M. (2007). Effectiveness of cognitive behavioural group therapy for social phobia in a community clinic: A benchmarking study. *Behaviour Research and Therapy*, 45, 3030-3040.

- McEvoy, P. M., & Nathan, P. (2007). Effectiveness of cognitive behaviour therapy for diagnostically heterogeneous groups: A benchmarking study. *Journal of Consulting and Clinical Psychology, 75*, 344-350.
- Mitte, K. (2005). A meta-analysis of the efficacy of psycho- and pharmacotherapy in panic disorder with and without agoraphobia. *Journal of Affective Disorders, 88*, 27-45.
- Norton, P. J. (2008). An open trial of a transdiagnostic cognitive-behavioral group therapy for anxiety disorder. *Behavior Therapy, 39*, 242-250.
- Norton, P. J., Hayes, S. A., & Hope, D. A. (2004). Effects of a transdiagnostic group treatment for anxiety on secondary depression. *Depression and Anxiety, 20*, 198-202.
- Norton, P. J., & Price, E. C. (2007). A meta-analytic review of adult cognitive-behavioral treatment outcome across the anxiety disorders. *Journal of Nervous and Mental Disease, 195*, 521-531.
- Oei, T. P. S., & Dingle, G. (2008). The effectiveness of group cognitive behaviour therapy for unipolar depressive disorders. *Journal of Affective Disorders, 107*, 5-21.
- Oei, T. P. S., Llamas, M., & Devilly, G. J. (1999). The efficacy and cognitive processes of cognitive-behaviour therapy in the treatment of panic disorder with agoraphobia. *Behavioural and Cognitive Psychotherapy, 27*, 63-88.
- Olatunji, B. O., Cisler, J. M., & Tolin, D. F. (2007). Quality of life in the anxiety disorders: A meta-analytic review. *Clinical Psychology Review, 27*, 572-581.
- Persons, J. B. (1991). Psychotherapy outcome studies do accurately represent current models of psychotherapy: A proposed remedy. *American Psychologist, 46*, 99-106.

- Persons, J. B., Bostrom, A., & Bertagnolli, A. (1999). Results of randomized controlled trials of cognitive therapy for depression generalize to private practice. *Cognitive Therapy and Research*, 23, 535-548.
- Rosenberg, N. K., & Hougaard, E. (2005). Cognitive-behavioural group treatment of panic disorder and agoraphobia in a psychiatric setting: A naturalistic study of effectiveness. *Nordic Journal of Psychiatry*, 59, 198-204.
- Stuart, G. L., Treat, T. A., & Wade, W. A. (2000). Effectiveness of an empirically based treatment for panic disorder delivered in a service clinic setting: 1-year follow up. *Journal of Consulting and Clinical Psychology*, 68, 506-512.
- Taylor, S. (1996). Meta-analysis of cognitive-behavioral treatments for social phobia. *Journal of Behavior Therapy and Experimental Psychiatry*, 27, 1-9.
- Tucker, M., & Oei, T. P. S. (2007). Is group more cost-effective than individual cognitive behaviour therapy: The evidence is not solid yet. *Behavioural and Cognitive Psychotherapy*, 35, 77-91.
- van Balkom, A. J. L. M., van Boeijin, C. A., Boeke, A. J., van Oppen, P., Kempe, P. T., & van Dyck, R. (2008). Comorbid depression, but not comorbid anxiety disorders, predicts poorer outcome in anxiety disorders. *Depression and Anxiety*, 25, 408-415.
- Wardle, J. (1990). Behavior therapy and benzodiazepines: Allies or antagonists? *British Journal of Psychiatry*, 156, 163-168.
- Weisz, J. R., Donenberg, G. R., Han, S. S., & Weiss, B. (1995). Bridging the gap between laboratory and clinic in child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology*, 63, 688-701.

- Westbrook, D., & Hill, L. (1998). The long-term outcome of cognitive behaviour therapy for adults in routine clinical practice. *Behaviour Research and Therapy*, 36, 635-643.
- Westbrook, D., & Kirk, J. (2005). The clinical effectiveness of cognitive behaviour therapy: Outcome for a large sample of adults treated in routine practice. *Behaviour Research and Therapy*, 43, 1243-1261.
- Zung, W. W. K. (1965). A self-rating depression scale. *Archives of General Psychiatry*, 12, 63-70.

Footnotes

¹One reviewer suggested that our results may be enhanced through the use of alternative analysis techniques such as principal components analysis (PCA) to reduce the number of outcome measures. Unfortunately, in our dataset the large amount of missing data precluded meaningful use of any such strategy. Examination of our dataset showed that only 80 participants had complete data on all outcome measures. Conducting a PCA with such a small subsample is dubious, and the extent to which the findings from these 80 participants could be meaningfully compared with the overall cohort is questionable. As such we elected not to conduct a PCA but to present the results in the form gathered from the original questionnaires.

Table 1

Comparison Between Treatment Completers and Non-Completers on Pre-Treatment Variables

Variable	α	Non-Completers	Completers	Comparison
Age		$M = 40.76, SD = 12.19, n = 120$	$M = 43.55, SD = 12.05, n = 176$	$t(294) = 1.95$
Beck Anxiety Inventory	.92	$M = 25.92, SD = 13.85, n = 125$	$M = 25.60, SD = 12.81, n = 181$	$t(304) = 0.21$
Depression Anxiety Stress Scale - Stress	.87	$M = 14.63, SD = 5.76, n = 81$	$M = 13.45, SD = 5.11, n = 134$	$t(213) = 1.55$
Mobility Inventory – Accompanied	.97	$M = 1.85, SD = 0.84, n = 109$	$M = 1.76, SD = 0.69, n = 152$	$t(259) = 1.04$
Mobility Inventory – Alone	.96	$M = 2.37, SD = 1.09, n = 106$	$M = 2.28, SD = 1.00, n = 146$	$t(250) = 0.70$
Fear Questionnaire – Total	.73	$M = 39.17, SD = 21.16, n = 122$	$M = 38.05, SD = 22.20, n = 176$	$t(296) = 0.41$
Panic and Agoraphobia Scale	.88	$M = 1.54, SD = 0.87, n = 100$	$M = 1.50, SD = 0.84, n = 157$	$t(255) = 0.39$
Zung Self-Rating Scale for Depression	.88	$M = 54.26, SD = 6.54, n = 82$	$M = 52.75, SD = 6.35, n = 134$	$t(214) = 1.67$
Catastrophic Cognitions Questionnaire	.95	$M = 64.40, SD = 17.78, n = 123$	$M = 61.25, SD = 16.84, n = 181$	$t(302) = 1.56$
Quality of Life Inventory	.85	$M = 62.79, SD = 15.37, n = 80$	$M = 65.28, SD = 15.85, n = 129$	$t(207) = 1.12$
Satisfaction with Life Scale	.61	$M = 13.54, SD = 6.42, n = 123$	$M = 15.38, SD = 7.79, n = 180$	$t(301) = 2.15^*$

* $p < .05$

Table 2

Significance Tests and Effect Sizes of Treatment on Outcome Variables (Complete Sample)

Outcome Measure	Pre-Treatment	Post-Treatment	Significance and Effect Size
Beck Anxiety Inventory	$M = 25.60, SD = 12.81$	$M = 17.39, SD = 13.06$	$t(180) = 10.03^{**}, d = 0.64$
Depression Anxiety Stress Scale – Stress	$M = 13.63, SD = 5.18$	$M = 10.08, SD = 5.01$	$t(151) = 8.17^{**}, d = 0.69$
Mobility Inventory – Accompanied	$M = 1.77, SD = 0.64$	$M = 1.60, SD = 0.80$	$t(145) = 3.78^{**}, d = 0.27$
Mobility Inventory – Alone	$M = 2.33, SD = 0.99$	$M = 2.11, SD = 0.96$	$t(135) = 3.95^{**}, d = 0.22$
Fear Questionnaire	$M = 39.17, SD = 22.57$	$M = 31.79, SD = 21.57$	$t(211) = 7.07^{**}, d = 0.33$
Panic and Agoraphobia Scale	$M = 1.52, SD = 0.83$	$M = 1.06, SD = 0.87$	$t(154) = 8.80^{**}, d = 0.55$
Zung Self-Rating Scale for Depression	$M = 53.04, SD = 6.55$	$M = 52.24, SD = 6.14$	$t(153) = 1.60, d = 0.12$
Catastrophic Cogn. Questionnaire	$M = 61.93, SD = 16.81$	$M = 53.34, SD = 18.47$	$t(214) = 8.75^{**}, d = 0.51$
Quality of Life Inventory	$M = 64.73, SD = 16.11$	$M = 70.15, SD = 16.58$	$t(148) = 6.50^{**}, d = 0.34$
Satisfaction with Life Scale	$M = 14.99, SD = 7.62$	$M = 16.85, SD = 7.76$	$t(213) = 4.25^{**}, d = 0.24$

$^{**}p < .01$

Table 3

Significance Tests and Effect Sizes of Treatment on Outcome Variables (BAI ≥ 20)

Outcome Measure	Pre-Treatment	Post-Treatment	Significance and Effect Size
Beck Anxiety Inventory	$M = 33.28, SD = 9.17$	$M = 22.51, SD = 13.22$	$t(114) = 9.31^{**}, d = 1.17$
Depression Anxiety Stress Scale – Stress	$M = 15.56, SD = 4.32$	$M = 11.45, SD = 5.03$	$t(93) = 7.37^{**}, d = 0.95$
Mobility Inventory – Accompanied	$M = 1.92, SD = 0.70$	$M = 1.71, SD = 0.93$	$t(83) = 3.21^{**}, d = 0.30$
Mobility Inventory – Alone	$M = 2.58, SD = 1.06$	$M = 2.29, SD = 1.05$	$t(76) = 3.70^{**}, d = 0.27$
Fear Questionnaire – Total	$M = 45.28, SD = 23.33$	$M = 36.21, SD = 22.84$	$t(119) = 5.90^{**}, d = 0.39$
Panic and Agoraphobia Scale	$M = 1.86, SD = 0.72$	$M = 1.31, SD = 0.90$	$t(100) = 7.86^{**}, d = 0.76$
Zung Self-Rating Scale for Depression	$M = 52.10, SD = 6.54$	$M = 51.17, SD = 5.72$	$t(95) = 1.45, d = 0.06$
Catastrophic Cogn. Questionnaire – Total	$M = 66.93, SD = 15.01$	$M = 56.34, SD = 19.53$	$t(124) = 7.50^{**}, d = 0.71$
Quality of Life Inventory	$M = 61.25, SD = 15.29$	$M = 66.89, SD = 16.08$	$t(90) = 5.21^{**}, d = 0.37$
Satisfaction with Life Scale	$M = 13.71, SD = 6.75$	$M = 15.40, SD = 7.00$	$t(123) = 2.86^{**}, d = 0.25$

$^{**}p < .01$

Table 4

Benchmarking Comparisons with Earlier Effectiveness Studies

Study	N	Diagnosis	Measure	Effect Size
Westbrook & Kirk (2005)	77	GAD	BAI	0.7
	76	Social Phobia	BAI	0.4
	95	Specific Phobia	BAI	0.5
	39	Agoraphobia	BAI	0.8
	46	Panic Disorder	BAI	0.8
McEvoy & Nathan (2007)	30	Heterogenous Anxiety	BAI	0.4
	75	Anxiety and Depression	BAI	0.4
Rosenberg & Hougaard (2005)	53	Panic and Agoraphobia	STAI-S	0.58
	53	Panic and Agoraphobia	STAI-T	0.48
Current Study	396	Heterogenous Anxiety	BAI	0.64

Figure Captions

Figure 1. Normative comparison.