Treatment of obsessive–compulsive disorder: Cognitive behavior therapy vs. exposure and response prevention

Maureen L. Whittal*, Dana S. Thordarson, Peter D. McLean

Department of Psychiatry, University of British Columbia, Vancouver, BC, Canada V6T 2B5

Received 9 April 2003; received in revised form 28 October 2004; accepted 4 November 2004

Abstract

The efficacy of contemporary cognitive therapy for obsessive–compulsive disorder (OCD) has only recently been investigated. The current study compares exposure and response prevention (ERP) and cognitive behavior therapy (CBT) delivered in an individual format. Participants were randomly assigned to the 12 consecutive-week CBT or ERP treatment. Based on 59 treatment completers, there was no significant difference in YBOCS scores between CBT and ERP at post-treatment or at 3-month follow-up. A higher percentage of CBT participants obtained recovered status at post-treatment (67%) and at follow-up (76%), compared to ERP participants (59% and 58%, respectively), but the difference was not significant. Effect sizes (ESs) were used to compare the results of the current study with a previous study conducted at our center that utilized group CBT and ERP treatments, as well as other controlled trials that have compared CBT and ERP. The significance of these results is discussed and a comparison is made with the existing literature.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: Obsessive–compulsive disorder; Cognitive behavioral treatment; Exposure and response prevention

*Corresponding author. Anxiety Disorders Clinic, UBC Hospital, 2211 Wesbrook Mall, Vancouver, B.C., Canada V6T 2B5. Tel.: +1 604 822 1789; fax: +1 604 822 7452.
E-mail address: whittal@interchange.ubc.ca (M.L. Whittal).
Introduction

Since the pioneering work of Meyer (1966), exposure and response prevention (ERP) has been firmly established as the psychosocial treatment of choice for obsessive–compulsive disorder (OCD). In a review of the behavioral literature, Baer and Minichiello (1998) reported that following a course of ERP, 60–70% of treatment completers (with observable compulsive behaviors) were “much improved”. However, despite such promising results, ERP is not a panacea. There are those participants who do not improve with treatment (particularly people with primary obsessions; see Rachman, 1983), other participants find ERP difficult to tolerate, and the majority of treatment completers are left with residual symptoms. Taking into account treatment refusers/drop outs/failures, Stanley and Turner (1995) indicated that acute treatment success declines to 63% of those who sought treatment and a further decline to 55% at long-term follow-up (1–6 years following acute treatment).

Cognitive treatment for OCD was developed, in part, to target distorted cognitions conceptualized as potential maintaining factors that remain unaddressed with behavioral treatment and to provide a less anxiety-provoking alternative to ERP, which would hopefully result in lower drop out rates. Although recognition of the importance of interpretation of intrusive thoughts dates back over 25 years (e.g., Carr, 1974; McFall & Wollersheim, 1979), Salkovskis (1985) put forward the contemporary cognitive model of OCD. The cornerstone of this model is the interpretation of common and unwanted intrusive thoughts. Rachman and de Silva (1978) reported that greater than 90% of a community sample experienced intrusive unwanted thoughts. Salkovskis and Harrison (1984) replicated this finding with an analogue sample. As such, it is not the intrusive thoughts that are a problem, but the interpretation that follows. Salkovskis (1985, 1996) asserted that faulty interpretations or appraisals of responsibility result in anxiety and the urge to neutralize. For example, the appraisal of “It will be my fault if something bad happens” following an intrusive doubt of the stove being on, leads to emotional distress and the urge to neutralize (checking in this example). Rachman (1997, 1998) has suggested appraisals that indicate the person is “mad, bad, or dangerous” will result in distress and the urge to neutralize (e.g., the appraisal of “I’m a sicko” following an unwanted aggressive thought). The obsessive–compulsive cognitions working group (OCCWG, 1997) have suggested other potential appraisals of intrusive thoughts that may be important in OCD (e.g., overestimations of threat, need to control thoughts, and overimportance of thoughts) in addition to inflated responsibility.

To date, there have been relatively few published trials of the efficacy of contemporary cognitive treatment of OCD.¹ In a series of individual treatment studies, Emmelkamp and colleagues employed early cognitive treatments (e.g., Emmelkamp & Beens, 1991; Emmelkamp, van der Helm, van Zanten, & Plochg, 1980; Emmelkamp, Visser, & Hoekstra, 1988) but they did not target the appraisals suggested by the current cognitive theory. Freeston et al. (1997), van Oppen et al. (1995) and Cottraux et al. (2001) have published the only controlled trials of

¹To avoid confusion between the use of terms CBT, CT, and behavioral treatments, ‘contemporary cognitive treatment refers to therapy that emphasizes cognitive challenging of distorted thoughts/appraisals and hypothesis testing with no emphasis on response prevention and tolerating or habituation to anxiety. Throughout the remainder of the manuscript we will refer to this treatment as CBT in keeping with our previous published paper (McLean et al., 2001).
individual contemporary CBT for OCD, with the latter two studies comparing the efficacy of CBT and ERP. van Oppen et al. (1995) compared cognitive treatment to in vivo exposure and response prevention over 16 weekly sessions that were 45 min in duration. Both groups achieved significant declines in YBOCS scores from pre- to post-treatment, although a higher percentage of cognitively treated patients (39%) achieved “recovered” status compared to ERP (17%) patients. However, as noted by Steketee, Frost, Rhéaume, and Wilhelm (1998), the mean post-treatment YBOCS for the ERP group was higher (17.3) than is typical, which may have made the cognitive treatment look better by comparison.

Cottraux et al. (2001) compared 65 patients randomly assigned to either individual cognitive therapy (e.g., Beck, Emery, & Greenberg, 1985; Salkovskis, 1985) or intensive behavior therapy over 16 weeks. Unfortunately, those participants who had Hamilton depression scores above 20 were excluded, which may limit generalizability of the findings. At post-treatment and one-year follow-up, there were no significant differences between the groups. Although the pre-treatment YBOCS scores were higher (approximately 28) than most published studies, the post-treatment YBOCS were also higher than normal at approximately 16.

The present study follows that of a randomized controlled trial conducted earlier at our center (McLean et al., 2001), which compared group ERP and group CBT. McLean et al. (2001) reported at post-treatment while controlling for medication use, the ERP group had significantly lower YBOCS scores than the CBT group. This difference was maintained at 3-month follow-up. Moreover, at 3-month follow-up a significantly higher percentage of ERP treatment completers (44%) reached recovered status compared to CBT completers (13%). McLean et al. (2001) suggested that ERP was easily adapted to be delivered in a group format whereas the idiosyncratic nature of the appraisals in OCD would likely favor individual CBT over group CBT. The purpose of the present study is to investigate the efficacy of CBT delivered individually given the relatively poor performance of group CBT for OCD (McLean et al., 2001).

The present study allows a comparison to McLean et al. (2001) study as a new cohort of OCD participants were randomly assigned to individual CBT or ERP. The results of the individual comparison of CBT and ERP will be examined at post-treatment and at 3-month follow-up. Due to the sequential delivery of group and individual treatment, the results of the two studies cannot be combined. However, using effect sizes (ES), the individual treatment outcomes will be compared to the group outcomes of McLean et al. (2001) as the treatment manuals, procedures, and lead therapists were the same. Additionally, we will also compare the ESs of other controlled studies that compared the efficacy of CBT and ERP for OCD.

Method

Participants

The 83 participants who were offered and accepted individual treatment were primarily (i.e., 74%) referred by physicians. Eighteen percent were self-referred and the remainder were referred by other health professionals. Inclusion criteria were a primary diagnosis of OCD according to the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV, American Psychiatric Association, 1994), an age range of 18–65 yr, fluency in written and spoken English, and
minimum symptom duration of one year. Exclusion criteria included active thought disorder or uncontrolled bipolar disorder, mental retardation, organic mental disorder, initiation or change in medication three months prior to their intake assessment, and unwillingness to refrain from making any changes in medication during treatment. Concurrent psychological treatment for any other Axis I or II disorder also resulted in exclusion, with the exception of marital therapy or supportive therapy for depression.

**Design**

Participants were randomly assigned to receive either CBT or ERP. There was no waiting list condition as previous studies at our center (e.g., McLean et al., 2001) indicated that OCD symptom severity remained stable during the 3-month delay. All participants were assessed pre-treatment, post-treatment, and three months post-treatment.

**Procedure**

Prior to intake assessment, participants were screened by telephone. Intake assessment consisted of the Structured Clinical Interview for DSM-IV (SCID; First, Spitzer, Gibbon, & Williams, 1996) to determine diagnoses for current Axis I disorders and the Yale-Brown Obsessive–Compulsive Scale (YBOCS; Goodman et al., 1989) to determine the severity of the OCD symptoms.

If inclusion criteria were met during the initial interview the participant was randomly assigned to ERP or CBT and was seen by the treating clinician for a second assessment. These assessments averaged approximately 3 h and included a review of OCD symptomatology, social history, identification of pre-disposing, precipitating, and maintaining factors, and an explanation of the model that was to guide their treatment. For participants who were treated with ERP, an exposure hierarchy was developed. Following this second assessment, a case-based formulation was written and was discussed in weekly team meetings.

Participants completed a pre-treatment questionnaire packet between the two assessment sessions. The SCID, YBOCS, and questionnaire measures were completed again following treatment and three months later. The assessors were blind to treatment type. With few exceptions, the same assessor was used at pre-, post-, and follow-up. Prior to each therapy session, subjects completed a self-report YBOCS (Baer, Brown-Beasley, Sorce, & Henriques, 1993) and the frequency section of the Responsibility Interpretations Questionnaire (Salkovskis et al., 2000). After the second treatment session participants completed the Reaction to Treatment Questionnaire (Borkovec & Nau, 1972) to assess treatment expectations. Therapists completed a homework compliance form (Primakoff, Epstein, & Covi, 1986) from sessions 2 to 12.

**Measures**

1. **The Yale-Brown Obsessive–Compulsive Scale (YBOCS; Goodman et al., 1989)**. The YBOCS is a semi-structured interview designed to identify the duration, control, and distress associated with obsessive and compulsive behaviors. Items 1–5 are summed for a subscale score on obsessions and items 6–10 are summed for a subscale score on compulsions. In the current
study, the total score was used (sum of items 1–10). Each item is rated from 0 (none) to 4 (extreme); thus scores on the YBOCS range from 0 to 40. Taylor (1998) reported that the interrater reliability varies between .80 and .99. Two-week test–retest reliability of the YBOCS ranges between .81 and .97. Due to the high correlation with measures of depression and anxiety, discriminant validity is poor. For example, Taylor (1998) reported correlations of .53–.91 between YBOCS and Hamilton Depression Scale and .47–.85 with the Hamilton Anxiety Scale. The clinician administered version was used at pre-, post-, and follow-up. Participants completed the 10-item self-report version (Baer et al., 1993) each week of treatment. Steketee, Frost, and Bogart (1996) reported 1-week test–retest reliability from .82 to .88 and discriminates between OCD and non-OCD samples.

2. **Beck Depression Inventory** (BDI; Beck & Steer, 1987). The BDI is a 21-item questionnaire to rate the severity of depression. Items are rated from 0 to 3 and scores range from 0 to 63. It is a well-established measure and its psychometric properties are excellent.

3. **Obsessional Belief Scale** (OBQ; Obsessive–Compulsive Cognitions Working Group [OCCWG]) is a 44-item questionnaire designed to measure the strength regarding the beliefs thought to be related to obsessions. Items are rated from 1 to 7 and include 3 factor-analytically derived subscales: responsibility/harm, perfection/certainty, and importance/control. Internal consistency is high for the three subscales and it demonstrates good known groups validity in clinical and non-clinical participants \( r \) of .48 and .45, respectively. As this is a recent revision of the original OBQ, participants completed the longer version that was then scored in accordance with the 44-item version.

4. **Interpretations of Intrusions Inventory** (III; Frost & Steketee, 2002 (Appendix B)) is a 31-item questionnaire designed to assess the appraisals of unwanted intrusive thoughts. A recent factor analysis (OCCWG) revealed a single factor as opposed to the original three rationally derived subscales. Internal consistency for the III total score was high and the OCD sample scored significantly higher than anxious controls.

5. **Reaction to Treatment Scale** (Borkovec & Nau, 1972) is a 4-item scale designed to assess the degree to which participants expect the treatment to be successful. Each item is rated on a 0–10 scale. Thus scores range from 0 to 40. The reaction to treatment scale has been used in many treatment outcome studies and has well-established psychometric properties.

6. **Homework compliance** (Primakoff et al., 1986) is a 6-item scale designed to assess the degree of homework compliance. Items range from (1) the patient did not attempt the assigned homework to (6) the patient did more of the assigned homework than was requested. The items on this scale have not been subject to psychometric evaluation.

**Therapists**

Licensed clinical psychologists experienced in treating OCD or psychology interns conducted the treatment. The interns treated 11 individual patients. MLW supervised the psychology interns by audiotape review or in cotherapy. A randomly selected portion of individual patients’ sessions were audiotaped and rated by a peer reviewer for adherence to the treatment manual (e.g., did the therapists refrain from using cognitive methods in ERP). Ratings were based on a 0–10 scale and corresponded to pre-established guidelines. The average adherence to CBT and ERP treatments was 8.7, with a range of 6–10. Mean adherence to the cognitive protocol was 8.62 (range 7.8–10).
Mean adherence to the behavioral protocol was 8.81 (range 6–10). These values indicate that adherence to the behavioral and cognitive–behavioral protocols were generally high.

_Treatments_

Individual treatment sessions were 50–60 min in duration and were delivered over 12 consecutive weeks.

**ERP.** The manual used in the current study was adapted from the one used by Van Noppen, Steketee, McCorkle, and Pato (1997). There were no cognitive elements in this treatment. Success was framed as habituation to a triggering stimulus and it was explained that obsessions were maintained through the negative reinforcement experienced following the completion of a compulsion. If patients initiated a discussion of cognitive issues (e.g., overestimation of danger), therapists did not respond and attempted to move the discussion back to more behavioral topics. For more information regarding the ERP treatment, refer to McLean et al. (2001).

**CBT.** The treatment manual for this condition was created at our center and was based upon the work of Salkovskis (1996), Freeston et al. (1996), and van Oppen and Arntz (1994). The treatment focused on challenging appraisals of intrusive thoughts (e.g., it will be my fault if something bad happens). Behavioral experiments were used to collect evidence for and against the alternate appraisals developed out of the cognitive challenging. Although these behavioral experiments could be considered as exposure, the purpose of them was to collect information for the alternate appraisal and not for habituation as in ERP. For a more complete explanation of the CBT used at our site and in the current study see Whittal and McLean (1999, 2002), Whittal, Rachman, and McLean (2002), and McLean et al. (2001).

_Results_

_Atrrion_

Eighty three participants met study criteria, were offered treatment, and accepted, at least tentatively. See Table 1 for a summary of attrition and outcome.

Three participants were excluded pre-treatment. Two potential participants were monozygotic twins who lived together, so we could not randomly assign them to different conditions without risking contamination of the treatments, and 1 participant was found to have a recent history of alcohol abuse.

Four participants were excluded from the study during treatment: 1 participant changed medications, 1 participant was hospitalized due to medical problems, 1 participant was found to have ego-syntonic homicidal thoughts, and therefore was considered too much of a safety risk to participate in the research protocol, and 1 participant was found to have severe depression with psychomotor retardation, as well as difficulty with written and spoken English. There was no significant difference between the treatment groups in the proportion of participants excluded from the study before or during treatment, $\chi^2 (1, N = 83) = 0.18, p = .67$. Participants who were
excluded from the study were not included in any of the subsequent analyses, leaving 76 participants who met study criteria.

Five participants dropped out of the study after their assessment but before the first session of treatment (i.e., treatment refusers). There was no significant relationship between refusing to participate and treatment condition, $\chi^2 (1, N = 76) = 0.27, p = .60). There was no significant difference between treatment refusers and non-refusers in education level, employment status, disability status, ethnicity, medication use, presence of a comorbid Axis I disorder, pre-treatment YBOCS, or age of onset (all $p > .05$). We were unable to determine whether treatment refusers were more likely to be depressed, because only 1 refuser returned the pre-treatment questionnaire package. All of the treatment refusers had been referred to the study by a physician or other health care provider; however, the association between referral source and refusing treatment was not statistically significant, $\chi^2 (1, N = 76) = 1.21, p = .27). A significantly higher proportion of men (4 out of 29) than women (1 out of 47) refused treatment, $\chi^2 (2, N = 76) = 3.97, p = .05$. All of the treatment refusers were single (5 out of 35 single people), resulting in a statistically significant association between marital status and treatment refusal $\chi^2 (1, N = 76) = 6.27, p < .05$. Treatment refusers were significantly younger ($M = 24.6$ yr) than non-refusers ($M = 34.6$ yr), $t(74) = 2.05, p = .04$, and refusers had OCD for less time ($M = 5.0$ yr vs. $M = 12.9$ yr), $t(14.4) = 4.34, p = .001$, $t$-test corrected for heterogeneous variances. Most of the patients who refused treatment reported doing so because they had too many work commitments or had difficulties obtaining transportation.

Not including the participants who were excluded from the study and participants who refused to participate, a total of 71 participants began treatment, attending at least one session of CBT

### Table 1

<table>
<thead>
<tr>
<th>Participant attrition and outcome</th>
<th>Total sample</th>
<th>Individual CBT</th>
<th>Individual ERP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Offered treatment and accepted</td>
<td>83</td>
<td>41</td>
<td>42</td>
</tr>
<tr>
<td>Excluded prior to treatment</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Excluded during treatment</td>
<td>4</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Dropped out prior to treatment</td>
<td>5</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Started treatment</td>
<td>71</td>
<td>34</td>
<td>37</td>
</tr>
<tr>
<td>Dropped out during treatment</td>
<td>12</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>Completed treatment &amp; post-treatment YBOCS</td>
<td>59</td>
<td>30</td>
<td>29</td>
</tr>
<tr>
<td>Met criteria for recovered status at post-treatment</td>
<td>37</td>
<td>20</td>
<td>17</td>
</tr>
<tr>
<td>Available at follow-up</td>
<td>53</td>
<td>29</td>
<td>24</td>
</tr>
<tr>
<td>Recovered at follow-up</td>
<td>36</td>
<td>22</td>
<td>14</td>
</tr>
</tbody>
</table>

**Notes.** To meet criteria for recovered status, participants had to show reliable change on the YBOCS total score (a decrease of at least 6 points), as well as have a final YBOCS total score below clinical range (less than 12). CBT = Cognitive-behavior therapy. ERP = Exposure and response prevention. YBOCS = Yale-Brown obsessive–compulsive scale total score.
(N = 34) or ERP (N = 37). Participants were considered to have completed treatment if they attended at least 7 sessions (i.e., more than half of the treatment program) and completed the post-treatment interview assessment; participants were considered to have dropped out of treatment if they stopped coming to sessions, expressed interest in terminating treatment before the final sessions, or were unavailable for the post-treatment interview assessment. Based on these criteria, 12 participants dropped out of treatment: 4 from CBT (12%) and 8 from ERP (22%), a non-significant difference, \( \chi^2 (1, N = 71) = 1.23, p = .27 \). Compared to participants who completed treatment, those who dropped out did not differ significantly in referral source, gender, education, disability status, marital status, ethnicity, medication use, presence of a comorbid Axis I disorder, age, age of onset, duration of OCD, or pre-treatment YBOCS (p’s > .05). Participants who dropped out differed in employment status; in particular, they were more likely to be students (4 out of 10 students dropped out) or unemployed but seeking work (5 out of 10 unemployed people dropped out). The association between employment status and dropping out was significant, \( \chi^2 (1, N = 71) = 17.09, p = .01 \). Participants who dropped out of treatment were more depressed, based on BDI scores at pre-treatment (dropouts N = 10, \( M = 25.70, SD = 3.95 \); completers N = 57, \( M = 17.72, SD = 10.09 \)), \( t (34) = 4.37, p < .001 \), t-test corrected for heterogeneous variances. Based on ratings available for only 6 of the 12 dropouts, they did not differ on treatment credibility. Reasons cited for dropping out of treatment included difficulty getting off work to come to sessions, serious illness, and family conflicts.

A total of 59 participants completed treatment.

**Demographics of treatment completers (N = 59)**

Participants (22 men, 37 women) had an average age of 35 years (range 18–65). The sample was predominantly European–Canadian (85%), with 12% Asian–Canadian, and 3% of other backgrounds. As a group, they were relatively well educated: 44% had a university degree, 27% had some post-secondary education, 20% had graduated from high school, and 8% had some high school education. A large proportion of the sample had never married (41%), 53% were married or cohabiting, and 7% were separated or divorced. Nearly half the participants were working full-time (49%), 15% were working part-time, 10% were students, 8% were unemployed, 8% were receiving disability benefits, 7% were full-time homemakers, and 2% were retired. Table 2 illustrates demographic characteristics by treatment type.

Of the 59 participants who completed treatment, 40 (68%) were currently using medication for their OCD. Forty-one percent were using a selective serotonin reuptake inhibitor (SSRI) alone, 17% were using multiple medications (typically an SSRI plus a benzodiazepine), 5% were using a tricyclic antidepressant alone, 3% were using only benzodiazepines, and 2% were using only other types of medication (e.g., mood stabilizers). Nearly half the participants (46%) had one or more secondary Axis I disorders, the most common of which were current major depressive disorder (25% of whole sample), generalized anxiety disorder (8%), and panic disorder with or without agoraphobia (8%). OCD spectrum disorders were uncommon: only 1 participant had Tourette’s disorder, and 1 participant had trichotillomania. A health care professional (e.g., primary care physician, psychiatrist, psychotherapist) referred 81% of participants and 19% were self-referred. The mean age of onset of OCD symptoms was 22 yr, and the mean duration of OCD symptoms was 13 yr.
For treatment completers \((N = 59)\), there were no significant differences between the CBT and ERP groups on demographic variables of gender, age, ethnicity, education, marital status, employment status, disability status, referral source, or medication use \((p's > .05)\). In addition,
there were no significant differences among the groups in age of onset of OCD symptoms, duration of OCD symptoms, YBOCS total scores prior to treatment, BDI scores prior to treatment, treatment credibility ratings, homework compliance, or the number of sessions attended (all \( p \)'s > .05).

Outcome for treatment completers: post-treatment and follow-up

Table 2 illustrates YBOCS means and standard deviations at pre-, post-, and follow-up. Within each treatment condition (CBT and ERP), there was a significant decrease in YBOCS scores from pre-treatment to post-treatment and from pre-treatment to follow-up (\( p \leq .001 \) for all paired-samples \( t \)-tests).

The two treatment groups (CBT and ERP) were compared at post-treatment using a one-way ANCOVA of YBOCS total scores, with pre-treatment YBOCS total scores as a covariate. There was no significant effect of treatment type at post-treatment; \( F(1,56) = 0.10, p = .76, \eta^2 = .002 \). The adjusted post-treatment mean for CBT was 10.22, compared to 10.81 for ERP.

YBOCS interview data at 3-month follow-up was available for 53 of the 59 treatment completers (90%). CBT and ERP were compared at follow-up using a one-way ANCOVA of YBOCS total scores, with pre-treatment YBOCS total scores as a covariate. There was no significant effect of treatment type at follow-up; \( F(1,50) = 0.89, p = .35, \eta^2 = .018 \). The adjusted follow-up treatment mean for CBT was 8.95, compared to 11.11 for ERP.

Intent-to-treat analysis using YBOCS self-report

The YBOCS self-report measure, which participants completed at each treatment session, enabled us to include 10 of the 12 participants who dropped out during treatment in an intent-to-treat analysis, using the self-report YBOCS scores at the last session attended as the outcome measure. (The other two participants had missing data for self-report YBOCS, having attended only one session and not completed questionnaires at that session.) YBOCS self-report scores were highly correlated with the interview YBOCS scores at pre-treatment (\( r = .53, p < .001, N = 69 \)) and at post-treatment (\( r = .90, p < .001, N = 59 \)).

First, the progress in treatment of completers and dropouts was compared with an ANCOVA of last session self-report YBOCS score, using first session self-report YBOCS score as a covariate. Not surprisingly, there was a significant effect of dropping out of treatment, \( F(1,66) = 9.43, p = .003, \eta^2 = .125 \). The adjusted post-treatment last self-report YBOCS means were 18.72 for treatment dropouts and 11.98 for treatment completers. Inspection of the unadjusted means revealed that treatment dropouts experienced little improvement (first YBOCS \( M = 26.20 \), last YBOCS \( M = 20.70 \)) whereas treatment completers improved substantially (first YBOCS \( M = 21.88 \), last YBOCS \( M = 11.64 \)). The mean number of sessions attended by participants who dropped out was 5.5; in contrast, treatment completers attended an average of 10.8 sessions.

An intent-to-treat analysis of treatment outcome comparing CBT with ERP was conducted, using a one-way ANCOVA of last self-report YBOCS score, with first self-report YBOCS score as covariate. There was no significant difference between the groups; \( F(1,66) = 0.14, p = .71, \eta^2 = .002 \).
The adjusted last session self-report YBOCS score was 13.26 for CBT compared with 12.66 for ERP.

Clinically significant change

In our previous paper (McLean et al., 2001), we used the method suggested by Jacobson and Truax (1991) for identifying participants who had experienced clinically significant change, based on a combination of (a) reliable reduction in YBOCS interview total scores and (b) a post-treatment YBOCS score outside the dysfunctional range. We chose to use the same criteria for reliable change and cut-off score as van Oppen et al. (1995), to enhance comparisons between their study and ours. We have continued to use the same criteria in the present study to identify recovered participants: a decrease in interview YBOCS of at least 6 points (a reliable change) and a final YBOCS score of 11 or less (a score below the clinical range). The number of participants classified as recovered in each condition is indicated in Table 1.

Of the participants receiving CBT, 67% were classified as recovered at post-treatment, compared to 59% of those receiving ERP, a non-significant difference, $\chi^2 (1, N = 59) = 0.41, p = .52$. Similar results were obtained at 3-month follow-up: 76% of the CBT group was classified as recovered, compared to 58% of the ERP group. Again, the association between treatment condition and recovered status was not statistically significant, $\chi^2 (1, N = 53) = 1.85, p = .17$.

Homework compliance

Therapist ratings of participant homework compliance were collected at every session, enabling us to examine whether there were treatment differences in homework compliance, and whether homework compliance affected treatment outcome. We calculated a mean homework compliance rating for each participant; sessions that were missed or sessions where therapists failed to rate homework compliance were not included in the mean score. Means and standard deviations of homework compliance ratings are presented in Table 2. There was no difference in homework compliance ratings between the ERP and CBT conditions.

However, homework compliance was a significant predictor of treatment outcome. For treatment completers, homework compliance was not correlated with pre-treatment interview YBOCS ($r = .11, p = .41, N = 58$; homework compliance ratings missing for one participant), but it was moderately correlated with post-treatment YBOCS ($r = -.34, p = .01, N = 58$); this relationship remained significant when controlling for pre-treatment YBOCS scores (partial $r = -.40, p = .002, N = 58$). The direction of the relationship indicates that high homework compliance was associated with lower post-treatment YBOCS scores.

Depression

Self-report ratings of depression (BDI scores) were collected at pre-treatment, post-treatment, and follow-up; descriptive statistics for BDI scores are provided in Table 2. Questionnaire data (including BDI, OBQ, and III scales) was available for 90% or more of participants for whom we had YBOCS interview data, with the exception of the ERP group at follow-up, where there was a 71% return rate. There was no significant difference in YBOCS or BDI scores at post-treatment
between participants who returned follow-up questionnaires and those who did not. However, due to the differential return rate, comparisons of ERP and CBT at follow-up based on questionnaire data must be considered tentative.

Within each treatment condition, there were significant decreases in BDI scores from pre- to post-treatment (CBT: $t(28) = 5.89, p < .001$; ERP: $t(25) = 3.78, p = .001$) and from pre-treatment to follow-up (CBT: $t(25) = 6.75, p < .001$; ERP: $t(15) = 3.10, p = .007$). ANCOVA of post-treatment BDI scores with pre-treatment BDI scores as covariate revealed no significant difference between CBT and ERP, $F(1,52) = 0.29, p = .59; \eta^2 = .006$. Nor was there a significant difference between the groups in depression scores at follow-up, $F(1,39) = 0.37, p = .56; \eta^2 = .009$.

Depression scores were significantly correlated with YBOCS scores at pre-treatment, $r = .44, N = 57, p = .001$, at post-treatment, $r = .57, N = 57, p < .001$, and at follow-up, $r = .53, N = 43, p = .001$. As might be expected, improvement in OCD symptoms was associated with improvement in depression; there were significant correlations between BDI and YBOCS residual gain scores at post-treatment, $r = .52, N = 55, p < .001$, and at follow-up, $r = .64, N = 42, p < .001$. However, depression at pre-treatment did not predict treatment outcome; there was no significant relationship between pre-treatment BDI scores and YBOCS residual gain scores at post-treatment, $r = .15, N = 57, p = .27$, or YBOCS residual gain scores at follow-up, $r = .09, N = 52, p = .64$.

**Belief change**

We used the OBQ subscales and the III to determine if CBT and ERP resulted in differential cognitive change. Descriptive statistics for these scales are presented in Table 2. A MANOVA revealed a significant difference in OBQ subscales and III at post-treatment between participants who returned follow-up questionnaires and participants who failed to return questionnaires, ($F(4,51) = 2.66, p < .05$). Participants who did not return their questionnaires at follow-up had significantly higher scores on the OBQ subscales and III at post-treatment compared to those who did return their questionnaires. As such, conclusions based upon follow-up questionnaire data should be considered tentative.

Within both treatment conditions, there were significant declines in the OBQ subscales and in the III from pre- to post-treatment, and from pre-treatment to follow-up, based on paired-sample $t$-tests (all $p < .001$). A series of univariate ANCOVAs were conducted to evaluate whether participants in CBT experienced greater cognitive change than participants in ERP, using post-treatment or follow-up OBQ subscale or III scores, with the corresponding pre-treatment scores as the covariate. An adjusted $\alpha$ of .006 (.05/8) was used to control for family-wise Type I error. There were no significant differences between the conditions on any of the OBQ subscales or the III scale at post-treatment or follow-up, covarying pre-treatment scores. Thus, it appeared that change in interpretations of thoughts and beliefs were similar in CBT and in ERP, despite such cognitive phenomena being the main focus of treatment in CBT.

Correlation coefficients between residual gain scores for the belief measures and the YBOCS interview scores indicated that belief change was moderately associated with symptom improvement across the whole sample. Focusing on pre-post residual gains, the correlation between YBOCS change and OBQ-Responsibility and Threat change was .52 ($p < .001, N = 55$),
for OBQ-Perfectionism and Certainty, $r = .52 \ (p < .001, N = 55)$, for OBQ-Importance and Control of Thoughts, $r = .34 \ (p = .01, N = 55)$, and for III, $r = .60 \ (p < .001, N = 54)$. Results for pre–follow-up residual gains were very similar but not reported here due to problems with missing data. In summary, it appears that cognitive change is associated with symptom improvement, but there was no difference between the treatment conditions in either symptom improvement or cognitive change.

**Standardized mean change effect sizes and comparison with previous studies**

Standardized mean change ES were calculated separately for each treatment condition using YBOCS interview scores at pre-treatment, post-treatment and at follow-up (see Table 3). ES were calculated as pre-treatment–post-treatment YBOCS score divided by pre-treatment standard deviation, as recommended by Becker (1988). Becker’s correction for small sample bias was also used. Unfortunately, according to Cumming and Finch (2001), it is not possible to estimate confidence intervals for ES as we have estimated them here, as they follow neither a central nor a non-central t distribution. Inspection of the effect sizes reveals that CBT participants improved

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Standardized mean change effect sizes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ES</td>
</tr>
<tr>
<td>Current study (individual)</td>
<td>CBT pre–post</td>
</tr>
<tr>
<td></td>
<td>ERP pre–post</td>
</tr>
<tr>
<td></td>
<td>CBT pre–follow-up</td>
</tr>
<tr>
<td></td>
<td>ERP pre–follow-up</td>
</tr>
<tr>
<td>McLean et al. (group)</td>
<td>CBT pre–post</td>
</tr>
<tr>
<td></td>
<td>ERP pre–post</td>
</tr>
<tr>
<td></td>
<td>CBT pre–follow-up</td>
</tr>
<tr>
<td></td>
<td>ERP pre–follow-up</td>
</tr>
<tr>
<td>Van Oppen et al. (individual)</td>
<td>CBT pre–post</td>
</tr>
<tr>
<td></td>
<td>ERP pre–post</td>
</tr>
<tr>
<td>Cottraux et al. (individual)</td>
<td>CBT pre–post</td>
</tr>
<tr>
<td></td>
<td>ERP pre–post</td>
</tr>
<tr>
<td></td>
<td>CBT pre–follow-up</td>
</tr>
<tr>
<td></td>
<td>ERP pre–follow-up</td>
</tr>
<tr>
<td>Freeston et al. (individual)</td>
<td>CBT pre–post</td>
</tr>
<tr>
<td></td>
<td>CBT pre–follow-up</td>
</tr>
</tbody>
</table>

**Notes.** Effect sizes based on interviewer YBOCS total scores. For Freeston et al., standard deviation and confidence intervals based on estimated pre–post and pre–follow-up correlation coefficients of $r = .4$ and .3, respectively. Formulae for standardized mean change ($g$), unbiased standardized mean change ($d$), and approximate standard deviation of $d$ ($s_d$) based on Becker (1988):

\[
g = \frac{M_{\text{post}} - M_{\text{pre}}}{\sqrt{s_{\text{pre}}^2}}
\]

$d$ is effect size corrected for small sample bias, $d = g \left(1 - \frac{3}{4n+1}\right)$. $s_d$ is approximate standard deviation of $d$, $s_d \approx \frac{2(1-r^2)}{n} + \frac{\hat{d}^2}{\frac{3}{4n+1}}$.
approximately 3 standard deviations from pre-treatment to follow-up, whereas ERP participants improved approximately 2 SDs.

To compare these results with previous findings in the literature, we calculated ESs and YBOCS percent change based on the results of McLean et al. (2001), van Oppen et al. (1995), Cottraux et al. (2001), and Freeston et al. (2000). Cottraux et al. reported their results in terms of gain scores; post-treatment and follow-up means and standard deviations were obtained via personal communication with Professor Cottraux. Examining the effectiveness of the various treatment conditions in these 5 studies (see Table 3), there is noticeable similarity in treatment response in most of the conditions across the studies. Participants tend to make an improvement of approximately 2–3 SDs regardless of treatment condition, with the exception of group CBT in the McLean et al. (2001) study, and individual ERP in van Oppen et al. In these two treatment arms, participants only improved approximately 1 SD; not surprisingly, these were the only studies out of the 5 to show significant difference between ERP and CBT for the treatment of OCD, although in different directions.

**Discussion**

In the present study the procedures of cognitive challenging of intrusive thoughts (i.e., CBT) and prolonged exposure (i.e., ERP) proved comparably effective as treatments for OCD delivered on an individual basis, in both treatment completers and intent-to-treat samples. This finding invites comparison to an earlier controlled trial investigating the relative efficacy of CBT versus ERP for OCD, when delivered in group format (McLean et al., 2001) in the same center, using the same treatment protocols. The earlier study found ERP to outperform CBT and we speculated that this result might be attributed to difficulty in engaging and challenging individual participant’s appraisals of intrusive thoughts in a group-based treatment. The advantage of group treatment for ERP is evident—it more easily allows other group members to observe one’s discomfort and behavioral engagement of feared stimuli, either directly, or indirectly, as in the case of self-reported exposure assignments (i.e., homework), and to cheer one’s efforts accordingly. In contrast, the discovery of errant cognitive appraisals and their replacement with more adaptive interpretations is a less tangible activity, and therefore one that is less amenable to group support. Further, in our group study, it was apparent that there was insufficient time for therapists using CBT to adequately identify and challenge each participant’s idiosyncratic cognitive distortions, raising the possibility that group CBT for OCD may under treat participants compared to individual therapy.

Given that the current study showed that cognitively focused treatment was at least as effective as exposure, when delivered individually, restores confidence in the efficacy of cognitive treatments and suggests the likelihood of a treatment type and modality interaction. Together, our two studies suggest that prolonged exposure can be delivered equally well in either individual or group modalities, whereas CBT is more effective when it is delivered individually. It suggests, for example, that CBT may become the psychological treatment of choice in cases where ERP is difficult to apply (e.g., primary obsessions).

Using ES calculations, we compared recent CBT and ERP studies in a mini meta-analysis. Given that the current study was conducted using the same treatment protocol and methodology,
except for modality of treatment, as McLean et al. (2001), we particularly interested in ESs of these two studies. With the exceptions of group CBT McLean et al. (2001) and the van Oppen et al. (1995) individual ERP, the ESs for the remaining treatment conditions demonstrated an improvement of 2–3 SD units from pre- to post-treatment, which was generally maintained through follow-up periods. The McLean et al. (2001) study reported a superiority of ERP over CBT, whereas van Oppen et al. (1995) reported the reverse. As others have noted (e.g., Steketee et al., 1998), van Oppen et al. (1995) used self-guided exposure that may not have been as effective as therapist-guided exposure. In fact, Abramowitz, Franklin, and Foa (2002) reported that self-controlled exposure resulted in significantly smaller pre–post reductions in OCD symptoms (27%) compared to therapist-assisted exposure (49%). Given that the van Oppen et al. (1995) study used a less effective form of ERP it is not surprising that they documented an advantage of CBT compared to ERP. Because ESs can be calculated differently, they are not all directly comparable. As a result, Table 3 lists pre- and post-treatment change scores on the YBOCS total measure, for each study, offering an alternative means to view treatment response.

Despite a differential treatment focus, prolonged exposure and CBT were associated with similar rates of cognitive change, as measured by the OBQ and the III, in keeping with the magnitude of symptom improvement. The question of whether OCD related cognitions are co-effects or causal in OCD symptom change remains unanswered. It is possible, for example, OCD cognitions for participants in exposure treatment become self-ameliorating, as a function of direct evidence that feared expectations never materialized. Had we debriefed participants about what they thought worked and how they interpreted exposure activities, we may have had some direction on this issue.

Given that the ERP delivered in the current study did not contain discussions of risk, uncertainty, and probability as is described in treatment manuals (e.g., Kozak & Foa, 1998), it is possible that the CBT looked better as it was compared to a substandard version of ERP. The average decline in the ERP group in the current study was 52% with an ES of 1.85. In a recent meta-analysis, Abramowitz et al. (2002) reported that the average symptom reduction for ERP was 48% (based upon 8 studies comparing ERP to a control, 7 of which were individual treatment studies) and the average ES was reported to be 1.50. Thus, although the current study may have offered a less complete ERP package, the decline in symptom severity was on par with studies that adopted a comprehensive ERP package. Despite the latter, it is possible that adding informal cognitive techniques to the current version of ERP may have resulted in even lower post-treatment YBOCS scores. However, as a dismantling study has not been completed to address the effect of informal cognitive techniques that are imbedded within ERP, the former possibility remains an empirical question.

In conclusion, both prolonged exposure and CBT protocols for OCD in this study yielded similar symptom reduction effects on the YBOCS measure of just over 50%, given 10–12 h of individual therapy. These results were maintained at the 3-month follow-up period. Approximately two-thirds of participants achieved and maintained clinical improvement status as a function of treatment, with a trend for more “recovered” participants in the CBT group. Course effects were not controlled for in this study and it is possible that a regression effect may account for the symptom improvement levels found in both groups. We think this unlikely however, given the documented stability in untreated OCD symptoms (e.g., McLean et al., 2001). Stronger treatment gains may be available by matching OCD sub-types to treatment orientation, including
provision for between session telephone on checks homework compliance, or treatment beyond 10–12 h.

Acknowledgements

We would like to acknowledge the dedication demonstrated by Maria Watson, M.A., and Lois-Jean Williams and the many undergraduate volunteers over the past several years. We would also like to acknowledge our many therapists who have included Quincy Young, Ph.D., Wendy Freeman, Ph.D., Mindy Smith, Ph.D., Sharon Agar, Ph.D., Lisa Brown, Ph.D., Melinda Losee, Ph.D., and Erica de Koning, Ph.D. We also appreciate the work of Sarah Newth Ph.D. who coded therapeutic alliance by listening to tapes. Last but not least, we would like to thank S. Rachman, Ph.D. for his comments on an earlier draft of the manuscript and Ralph Hakstian, Ph.D. for his statistical advice.

References


