A randomised controlled trial of acceptance-based cognitive behavioural therapy for command hallucinations in psychotic disorders

Frances Shawyer, John Farhall, Andrew Mackinnon, Tom Trauer, Eliza Sims, Kirk Ratcliff, Chris Larner, Neil Thomas, David Castle, Paul Mullen, David Copolov

A randomised controlled trial of acceptance-based cognitive behavioural therapy for command hallucinations in psychotic disorders.

Introduction

Command hallucinations are one of the most troubling symptoms of psychosis with numerous studies suggesting that they are associated with seriously destructive behaviours (see Braham, Trower, & Birchwood, 2004 for most recent review). Evidence has increasingly accumulated over the past 20 years that psychotic symptoms, which were often immutable even with antipsychotic medication, can be reduced or their impact ameliorated using Cognitive Behavioural Therapy (CBT) (Wykes, Steel, Everitt, & Tarrier, 2007). In the first randomised controlled trial (RCT) to apply these methods to command hallucinations, Trower et al. (2004) tested whether cognitive therapy directed towards modifying beliefs related to the power of the voice would reduce compliance with command hallucinations and increase resistance compared to a treatment-as-usual (TAU) control condition. They found a reduction in compliance in both groups at 6 months, which was significantly stronger in the treatment group and maintained at 12-months follow-up. The
results of this study are highly encouraging and support the efficacy of CBT for command hallucinations. However, using a TAU comparison condition means that the potentially influential contribution of non-specific treatment factors to the significant result cannot be evaluated (Lohr, Olatunji, Parker, & DeMaio, 2005; Lynch, Laws, & McKenna, 2010). In addition, the study focused on reducing compliance with severe command hallucinations, defined as harm to self or others, or major social transgressions; there was only limited examination of broader aspects of functioning.

Command hallucinations are of concern even when they are not associated with harmful content (Mackinnon, Copolov, & Trauer, 2004; Rogers, Gillis, Turner, & Frise-Smith, 1990; Romme, Honig, Noorthoorn, & Escher, 1992). Active efforts to resist command hallucinations may contribute to their malignancy, through increases in negative affect and greater engulfment with the symptom (see Shawyer, Farhall, Sims, & Copolov, 2005 for further review). Although concepts of resistance and non-compliance are sometimes conflated in relation to command hallucinations, resistance is most usefully defined as the degree to which a person wishes or attempts to oppose a command whereas non-compliance is the degree to which a person does not do what the command urges. It might be expected that measures of resistance and measures of non-compliance would largely overlap. However, evidence suggests that active attempts to resist command hallucinations may be of limited effectiveness in actually preventing compliance (Fox, Gray, & Lewis, 2004; Shawyer et al., 2008).

There is early evidence that methods from the so-called “third wave” of Behaviour Therapy (Hayes, 2004) that emphasize the mindful acceptance of experience may have broader benefits in relation to affect and function in psychosis. Acceptance and Commitment Therapy (ACT) is one of the most developed therapies from this newer genre. As the name suggests, ACT has two broad components. In the Acceptance component, “cognitive defusion” methods help the individual recognise and dispassionately observe symptoms as mental events, rather than judging their truthfulness and reacting to them. The Commitment component emphasizes the articulation of personal values and goals, and seeks to minimize the effects of symptoms on achieving those goals in behavioural terms. Bach and Hayes (2002) assessed the impact of a brief version of ACT in 80 inpatients with positive psychotic symptoms. They found that, compared with a TAU control group, participants in the ACT group believed less in the reality of their symptoms and had half the rate of rehospitalisation over a follow-up period of four months. This study was replicated with an enhanced ACT comparison group to control for amount of therapist contact (Gaudiano & Herbert, 2006). The results were comparable, and were evident across a wider range of outcomes, including distress and social functioning.

Following several decades of patchy and sometimes cautionary reports relating to meditation and psychosis, Chadwick, Newman, Taylor, and Abba (2005) published the first study to apply mindfulness meditation directly to psychotic symptoms. Mindfulness has been defined as “the awareness that emerges through paying attention on purpose, in the present moment, and non-judgmentally to the unfolding of experience moment by moment” (Kabat-Zinn, 2003, p. 145). Regular practice of mindfulness is thought to reduce reactivity to thoughts and emotions and increase self-awareness, self-acceptance and the capacity to make adaptive choices (Surawy, Roberts, & Silver, 2005). Taking into account the cautions contained in the earlier literature, Chadwick et al. conducted specially adapted mindfulness groups for people with psychosis focussing on helping participants to establish a mindful relationship with unpleasant symptoms. Although uncontrolled and with small numbers (n = 10), promising findings were reported in relation to general clinical functioning.

In developing more adaptive responses to command hallucinations, we considered that the application of acceptance-based strategies as an alternative to both compliance and resistance could lead to better outcomes not only in relation to compliance but also in relation to the broader impact of command hallucinations on quality of life and functioning as well as distress and general illness severity. On the basis of the promising findings to date, it has been suggested that including acceptance-based methods into CBT protocols could augment the effectiveness of CBT (Gaudiano, 2005). In the present paper we describe an RCT evaluating the efficacy of an Acceptance-Based CBT intervention for command hallucinations, “Treatment of Resistant Command Hallucinations” (TORCH) by comparison with a defined control intervention, Befriending. A waitlist condition was also included in order to be able to compare both treatments with TAU.

The primary hypothesis was that, compared to Befriending participants, TORCH participants would demonstrate reduced compliance with harmful command hallucinations, increased confidence in being able to resist obeying harmful commands, and increased confidence coping with commands. Secondary hypotheses were that TORCH participants would demonstrate lower illness severity and better functioning, reductions in distress and disruption to life and improvement in quality of life compared to controls. It was further hypothesised that any benefits associated with TORCH would be maintained at follow-up. In achieving these outcomes, we expected that changes would occur in a number of process variables for TORCH, but not Befriending, comprising: improved acceptance of command and auditory hallucinations and reduced involvement with them, and improved insight including lower conviction in key beliefs related to the reality and power of voices issuing commands. Finally, using the notions of recovery style developed McGlashan, Levy, and Carpenter (1975) we predicted that TORCH would promote a more integrated recovery style - that is, having a flexible and curious attitude to illness while Befriending would promote a sealing over response that seeks to isolate the illness from the sense of self and other aspects of life.

Method

Research design

The study was a prospective, single (rater)-blind, RCT using a part-parallel group comparison design involving the intervention, TORCH, and a comparison condition, Befriending. Both types of intervention were provided by all therapists involved in the trial. Participants continued their usual treatment during the course of their participation in the trial. A random sub-sample of participants was allocated to a 4-month waitlist prior to their treatment allocation. Assessments took place prior to commencing therapy (baseline), at the completion of therapy (endpoint), and at 6 months follow-up. Waitlist participants undertook an additional assessment prior to commencing the 4-month waiting period.

Power

With respect to the calculation of sample size, we determined a priori that a sample of 30 participants per treatment arm would enable moderate sized clinical effects (0.75 standard deviations difference in change) to be detected with power exceeding 80%, conservatively assuming a correlation of 0.50 between baseline and endpoint scores. The effect size chosen was based on the average effect size for CBT interventions reported by Rector and Beck (2001) at the time the study was planned (1.31 for positive symptom functioning) but with allowance for some improvement in outcome in the Befriending condition.
Recruitment and procedure

Recruitment took place from July 2003 to August 2005. Participants were recruited from a large number of public and private mental health services in metropolitan and regional areas of Victoria, Australia. Attempts were also made to recruit through forensic services but these were unsuccessful. Selection criteria included having a diagnosis of schizophrenia or related condition based on DSM-IV criteria, aged between 18 and 65 years and having experienced command hallucinations within the previous 6 months that caused distress or dysfunction despite treatment with antipsychotic medication at therapeutic doses. Exclusion criteria were: any neurological disorder that may affect cognitive function; insufficient conversational English for meaningful participation; current abuse of alcohol or drugs requiring specific clinical intervention; having a premorbid IQ of less than 70, and inability to give informed consent.

Patients referred to the project were interviewed to establish eligibility and obtain written consent. Eligible and consenting participants undertook a further interview for assessment with the outcome measures prior to randomisation. For approximately the first three-quarters of the recruitment period, participants were allocated to one of three groups, TORCH, Befriending or waitlist. To allow for waitlist participants to complete their therapy within the project’s time frame, any waitlist allocations made after this time were disregarded and the next treatment group allocated instead. To maximize the waitlist sample size, the ratio of allocation to treatment and waitlist allocation was 1:3. Randomization was by variable length blocks. Sequences of fixed length blocks were generated using the first random generator at www.randomization.com. Switching between blocks was determined by an additional random sequence generated by the same method. A similar procedure was used to generate assignments to treatments for participants initially assigned to the waitlist condition.

Assignment of participants to conditions was undertaken by the study statistician who worked independently of staff involved in the recruitment, assessment and management of participants in the study. Considerable efforts were made to ensure that the blindness of raters was maintained. Offices, data storage and travel logs of raters and therapists were kept separate and clinical staff and participants were regularly reminded not to divulge details of their therapy to the raters. Raters were asked to classify participants into a treatment condition before and after each assessment and to indicate their level of confidence. All breaches in blindness were recorded and addressed by changing the rater wherever possible.

Measures

Diagnosis and premorbid IQ

Diagnoses for psychotic disorders were established using appropriate sections of the Structured Clinical Interview for DSM-IV Axis 1 Disorders (First, Spitzer, Gibbon, & Williams, 1997). The National Adult Reading Test-Revised (NART-R - Crawford, 1990) was administered as a test of premorbid IQ.

Primary outcome measures

We assessed degree of compliance with harmful command hallucinations by adapting an interview schedule developed from an earlier study (Shawyer et al., 2008). Compliance was rated by the assessor on a scale of 0–7 (non-compliance – full compliance) for the maximal instance of compliance with the most recently experienced harmful commands. Confidence to resist obeying harmful commands and confidence in coping with general commands were self-rated using a scale of 0–100.

Secondary outcome measures

Illness severity and its effect on functioning was measured using the Positive and Negative Syndrome Scale (PANSS - Kay, 1991) and the Modified Global Assessment of Functioning scale, DSM-IV version (Modified GAF - Hall, 1995a, 1995b). The PANSS is a comprehensive symptom rating scale which consists of seven items measuring positive symptoms, seven items measuring negative symptoms and 16 items measuring general symptoms relevant to the experience of schizophrenia. The Modified GAF measures illness severity on the basis of both psychopathology and functioning (including occupational, social, and psychological functioning) on a 0 to 100 scale, with a lower score indicating poorer functioning. Where possible, ratings for the Modified GAF were provided by a primary clinician. The research assistant provided the ratings when there was no primary clinician was available (14% of assessments).

Distress related to command hallucinations was assessed by asking participants to rate how upsetting they found their most recent experiences of command hallucinations using a fully anchored 5-point scale from “not at all” to “overwhelming/terrible” (from the Single Hallucination Episode Record (SHER) – Farhall, 2005). The general severity of auditory hallucinations in terms of distress (amount” + “intensity”) and disruption was assessed using the relevant items from the Psychotic Symptom Rating Scales (Auditory Hallucinations) (PSYRATS - Haddock, McCarron, Tarrier, & Faragher, 1999). The PYSRATS is a structured interview designed to elicit details of an individual’s auditory hallucinations across a range of physical and psychological dimensions.

Quality of life was assessed using two subscales from the Quality of Life Enjoyment and Satisfaction Questionnaire (Endicott, Nee, & Harrison, 1993): Subjective Feelings and General Activities. These two subscales were selected on the basis of being brief, reasonably comprehensive and relevant to areas of focus in therapy. The Subjective Feelings subscale measures how much of the time participants feel positive (e.g., satisfied with life, happy) and how much of the time they feel able to function in society (e.g., to communicate, travel, deal with life’s problems). The General Activities subscale measures degree of satisfaction with general activities of life such as work, social relationships and ability to function. Overall life satisfaction and enjoyment were also measured using fully anchored 5-point scales.

Subjective feedback

At the end of their post-therapy assessment, participants were asked to complete the Client Satisfaction Questionnaire (Attiksson & Zwick, 1982) to assess the acceptability and experience of therapy. Four additional questions were added to give participants an opportunity to provide more specific feedback about their experience of therapy. These included ratings of emotional response and problem improvement with respect to command hallucinations and qualitative feedback related to therapy gains and disappointments.

Process measures

Acceptance of voices was assessed using the Voices Acceptance and Action Scale (VAAS – Shawyer et al., 2007). The VAAS is 31-item self-report questionnaire that assesses acceptance-based attitudes and actions in relation to general auditory (Section A) and command hallucinations (Section B). This scale was found to have good internal and test-retest reliability together with evidence of construct validity (Shawyer et al., 2007). The psychometrically improved 9-item version of Section A, the VAAS-9 (Rateiff, 2010) was used in preference to the 12-item version of Section A in the VAAS-31.

We used several scales to assess degree of involvement with auditory command hallucinations. Preoccupation was assessed
using a fully anchored 5-point scale rating how much of the time
was spent thinking about or talking with the voice(s) issuing
commands (from the SHER - Farhall, 2005). Engagement and
resistance, two dimensions of involvement with auditory halluci-
ations that are associated with poorer outcome (Shawyer et al.,
2005), were assessed using subscales of the same name from the
Beliefs about the Voices Questionnaire-Revised (BAVQ-R -
Chadwick, Lees, & Birchwood, 2000). Instructions for the BAVQ-R
were adapted to apply to the most dominant voice issuing
commands.

Insight was measured using the Insight Scale (Birchwood et al.,
1994), an 8-item self-report questionnaire that assesses three
dimensions of insight: awareness of symptoms, awareness of
illness, and perceived need for treatment. The belief about the
reality of voices issuing commands was assessed by a fully
anchored 5-point scale from “I was convinced the voice was real”
to “I was convinced the voice came from my own mind” (from the
SHER - Farhall, 2005). The belief about the power of the voices
issuing commands was assessed using the Omnipotence subscale of
the BAVQ-R. Recovery style was assessed using the Recovery Style
Questionnaire (RSQ - Drayton, Birchwood, & Trower, 1998), a self-
report questionnaire consisting of 39 statements that measures
adaptation to illness in terms of integration and sealing over.

Antipsychotic medication
The daily dose of antipsychotic medication was recorded at each
assessment point and converted to chlorpromazine equivalents.

Treatment groups
Clinical staff were made aware of treatment allocation and kept
informed of treatment progress in both groups. No attempt was
made to standardise routine clinical care including medication.
The design of the delivery was to aim for 15 weekly sessions of
approximately 50 min duration plus two follow-up sessions in the
six months following completion of therapy. Therapy sessions were
audiotaped for the purposes of quality control.

TORCH
TORCH was manualized, flexibly implemented according to an
individualised case formulation and delivered by a team of five
psychologists trained in CBT, ACT and mindfulness. TORCH
comprised three engagement and assessment sessions followed by
12 sessions at weekly intervals. Core modules included belief
modification and acceptance-based interventions. Supporting
modules included motivational interviewing, personalised psy-
choeducation, enhancing self-efficacy, relapse prevention, coping,
assertion and termination. Homework exercises were given where
feasible.

Belief modification involved identifying and modifying the key
beliefs that “hook” clients into compliance and lead to distress such
as the power of the voice. Where applicable, work also focused on
beliefs about self and associated delusions. The main components of
the acceptance arm of treatment as applied in TORCH were based
on Bach and Hayes (2002) and included:

a) Cultivating the capacity to just notice voices and associated
thoughts rather than believe and act on them;
b) Accepting voices even though one may not like them;
c) Accomplishing valued goals in the context of ongoing voices.

Mindfulness exercises were commenced within the first few
sessions; practice at home was encouraged with the provision of
audiotaped instructions. Exercises were relatively brief (typically
around 10 min) and graded, starting with general skills before
applying to command hallucinations with the aim of cultivating
a non-judgemental, non-reactive response to them. Exercises and
discussion related to identifying values and goals also generally
took place in the early part of therapy. ACT exercises and metaphors
were flexibly applied to foster detached acceptance of voices,
defusion from content and the capacity to act independently from
them, in line with values and goals.

Befriending
In a study of this kind it is crucial to control for therapist contact
and treatment expectancy (Lynch et al., 2010). This trial followed
the studies of Sensky et al. (2000), Turkington, Kingdom, and Turner
(2000) and Jackson et al. (2007) in using Befriending as the
comparison intervention. Befriending is a fully manualised control
intervention (Bendall, Killackey, Jackson, & Gleson, 2003) that
provides the patient with the same amount of therapist engage-
ment and expectancy as CBT and has similar drop-out rates (Bendall
et al., 2006). Befriending involves a series of conversations that are
like conversations with a friendly social acquaintance. The sessions
focus on neutral topics of interest and enjoyment for the client,
such as hobbies, sports, and current affairs (Bendall et al., 2003).
An explicit avoidance of discussion of symptoms and problems (redi-
recting the conversation back to the treatment clinician if needed)
provides the rationale for treatment and is likely to contribute to
positive expectancy (Bendall et al., 2006; Bendall et al., 2003).
Although Befriending mainly involves social conversation, in cases
where it is difficult to engage the participant in ongoing dialogue
e.g., few interests, low volition), neutral activities may be under-
taken. Activities are considered neutral if they do not involve
exposure to feared situations.

Quality control
Therapists in the trial were trained according to the TORCH and
Befriending manuals and attended weekly to fortnightly peer
supervision with JF as primary facilitator. As a subjective appraisal
of therapist adherence to guidelines, therapists recorded the
content of each session using Session Record forms, structured to
reflect the content of the manuals.

An audit of a random sample of 31 audiotaaped therapy sessions
was also undertaken by an independent auditor who was blind to
all participant data and to the audiotaape randomization procedure.
Tape randomisation was stratified according to therapist, therapy
(TORCH or Befriending), stage of individual therapy (early: sessions
1–8; late: sessions 9–15) and therapist experience (early vs late
case). To assess the quality of TORCH sessions and confirm that
Befriending sessions did not include TORCH techniques, sessions
were rated using an adaptation of the Cognitive Therapy Scale for
Psychosis (CTS-Psy -Haddock et al., 2001). The overall format and
system of scoring was the same as the original scale; however, we
altered the content of six items to include reference to ACT inter-
ventions or command hallucinations. To assess the quality of
Befriending sessions and ensure that TORCH did not include
Befriending techniques, sessions were also rated according to the
Befriending Treatment Integrity Measure, which is described in the
Befriending manual (Bendall et al., 2003). This measure consists of
five items which reflect the central components of Befriending,
such as “therapist redirects from discussion about symptoms to
a neutral topic”. A clinical psychologist (NT) who was blind to
participant data and audiotaape selection procedure performed the
ratings and assigned each session to either the TORCH or
Befriending group. The tape auditor was very experienced in CBT
for psychosis and familiar with ACT. In addition, he received
approximately 7 h training on the measures and the TORCH and
Befriending treatments.

Please cite this article in press as: Shawyer, F., et al., A randomised controlled trial of acceptance-based cognitive behavioural therapy for
Therapeutic alliance during treatment was assessed using the 19-item Revised Helping Alliance Questionnaire (HAq-II) (Luborsky et al., 1996). This instrument was administered by the therapist at the conclusion of sessions 4 and 10. The HAq-II is particularly relevant as it assesses the quality of the alliance independent of symptom improvement. Total scores range from 19 to 114. In order to reduce the tendency to provide socially desirable answers, participants were instructed to return the completed form in a sealed envelope and were informed that their therapist would not see the form during the course of therapy.

Overall trial quality was independently rated using the Clinical Trial Assessment Measure (CTAM -Tarrier & Wykes, 2004), a scale to assess the quality of trials based on CONSORT guidelines. The rating was undertaken by one of the developers of this instrument.

Statistical analysis

Hypotheses were tested using the General Linear Model program (analysis of covariance) in the Statistical Package for the Social Science/Predictive Analytics Software Statistics for Windows, Version 18. The statistical model was treatment condition as the independent variable and endpoint/follow up measurement as the outcome variable with baseline measurement as the covariate, adjusting for baseline values — a potential issue given baseline group differences on some measures. An endpoint comparison was conducted to test whether the intervention was effective and a follow up comparison was conducted to test for maintenance of treatment effects or the delayed emergence of treatment effects. Hypotheses were evaluated based on the treatment condition term. Modelling our analysis on Cather et al. (2005) who had a similarly small sample size, we also conducted within-group analyses using paired t-tests, using separate variance t-tests where appropriate, and examined the magnitude of effect sizes for the between-group comparisons. In view of 1) the small number of dropouts and 2) the small sample size, no attempts were made to impute missing data values — comparisons were made on the basis of the available data only.

Results

Participants

Sample description

From a total of 230 referrals, 168 candidates were assessed for eligibility. The remaining 62 referrals declined initial contact or could not be contacted thus preventing assessment of their eligibility. Of the 81 candidates assessed as eligible, 44 consented to participate and proceeded to randomisation (see Fig. 1). Our recruitment rate of 53% of eligible candidates is similar to the rate of 55% reported by Trower et al. (2004).

Participants were mostly born in Australia (81%) and 24 (56%) were male. The mean age was 39 years (SD: 10). At baseline, the majority of participants were in receipt of a disability pension (61%) with 19% in part-time or full-time employment. Fifty-one percent of participants were living in independent accommodation, 23% were living with family and 26% were in supported accommodation. The most common primary diagnosis was schizophrenia (72%), with 21% having schizoaffective disorder and the remaining 7% having an affective psychosis. Mean score on the Modified GAF was 44.4 (SD: 11.5) indicating that, on average, participants experienced serious symptoms or impairment in several areas of functioning. Clinical and demographic characteristics for the TORCH and Befriending group were shown in Table 1. Compared to the Befriending group, those in the TORCH group were more likely to be male, experience more severe negative and general symptoms, have a history of less severe command hallucinations and be less distressed by recent commands. There were no other significant differences between groups at baseline. A group \times time repeated measures analysis of variance also showed no significant changes on average in dose of neuroleptic medication over the course of treatment [time: F(1,33) = 2.77, p = .11; group: F(1,33) = 0.01, p = .91]; group \times time: F(1,33) = 2.29, p = .14.

Participant flow and attrition

As shown in the CONSORT diagram (Fig. 1), 21 participants were allocated to TORCH and 22 to Befriending. An additional participant was randomized but contact from her psychiatrist the day after randomization indicated the candidate was ineligible as she was showing a delayed response to a change in medication. A sub-sample of 17 participants was randomized to a waitlist control prior to commencing therapy. Attrition was low with only three participants withdrawing from the study. Two participants withdrew from Befriending after five and eight sessions respectively. Both participants reported that intrusions or commands from auditory hallucinations contributed to the decision to withdraw. One of these participants also felt that his paranoia worsened as a result of the social contact. The third participant withdrew from TORCH after three sessions due to gaining work interstate. In addition to these withdrawals, two participants received fewer than 15 sessions. One TORCH participant received 12 sessions and one Befriending participant received 9 sessions. It appeared that in both cases, difficult and chaotic personal circumstances reduced the motivation and ability to attend appointments.

Quality control

Analysis of session records

Analysis of the TORCH and Befriending session record forms suggested that the content of therapy followed the respective manuals and overall approaches of both therapies. In terms of the main component of therapy sessions for TORCH, of the 295 sessions with available data, 42% involved acceptance-based interventions, 27% involved engagement and assessment, 19% involved belief modification, and 12% involved a therapy support module (over half of these related to future planning and review). With respect to Befriending, of the 297 sessions with available data, 91% involved neutral topics with a further 6% mainly involving neutral activities. The pattern of engagement for TORCH and Befriending was quite similar, with the modal rating for both being “good”. In only 8% of TORCH sessions and 7% of Befriending sessions did the therapist rate engagement as low or tenuous. There was no overall difference in the mean rating of engagement between groups [TORCH: 3.7 (SD: 0.52); Befriending: 3.9 (SD: 0.78); t(41) = −1.1, p = .29].

Independent audit of therapy sessions

The independent auditor correctly assigned all 31 taped sessions to the appropriate treatment condition (17 to TORCH and 14 to Befriending). With regard to treatment integrity, although many of the general items on the CTS-Psy were present in both forms of therapy, no TORCH session included a Befriending intervention and no Befriending session included a specific CBT or ACT intervention.

Treatment exposure

There were no significant differences between groups in the average length of sessions, [TORCH: 55.3 (SD: 5.0); Befriending: 53.1 (SD: 4.0); t(41) = 1.63, p = .11], total therapy time [TORCH: 13.1 (SD: 2.7); Befriending: 12.3 (SD: 2.7); t(41) = 1.06, p = .30], number of sessions [TORCH: 14.3 (SD: 2.7); Befriending: 14.4 (SD: 3.6); Mann–Whitney U test, p = .68], number of follow-ups [TORCH: 1.7

Please cite this article in press as: Shaway, F. et al., A randomised controlled trial of acceptance-based cognitive behavioural therapy for command hallucinations in psychotic disorders, Behaviour Research and Therapy (2012), doi:10.1016/j.brat.2011.11.007
Table 1

Demographic and clinical characteristics of TORCH and Befriending groups at baseline.

<table>
<thead>
<tr>
<th>TORCH (n = 17–21)</th>
<th>Befriending (n = 18–22)</th>
<th>Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years [Mean (SD)]</td>
<td>40.0 (8.5)</td>
<td>39.6 (11.4)</td>
</tr>
<tr>
<td>Sex (male/female)</td>
<td>15/6</td>
<td>9/13</td>
</tr>
<tr>
<td>NART IQ [Mean (SD)]</td>
<td>99.3 (14.3)</td>
<td>99.9 (13.4)</td>
</tr>
<tr>
<td>Years of education [Mean (SD)]</td>
<td>11.7 (2.6)</td>
<td>12.1 (2.1)</td>
</tr>
<tr>
<td>DSM-IV Diagnosis, n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>17</td>
<td>14</td>
</tr>
<tr>
<td>Schizoaffective Disorder</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Mood disorder with psychotic features</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Duration of illness, years [Mean (SD)]</td>
<td>14.2 (7.9)</td>
<td>15.2 (11.4)</td>
</tr>
<tr>
<td>Modified GAF score [Mean (SD)]</td>
<td>41.9 (11.9)</td>
<td>47.0 (10.7)</td>
</tr>
<tr>
<td>Chlorpromazine equivalent dose (mg) [Mean (SD)]</td>
<td>742.9 (388.7)</td>
<td>757.9 (560.6)</td>
</tr>
<tr>
<td>Involuntary status (yes/no)</td>
<td>2/18</td>
<td>3/18</td>
</tr>
<tr>
<td>PANSS score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive scale [Mean (SD)]</td>
<td>15.9 (4.0)</td>
<td>15.0 (3.6)</td>
</tr>
<tr>
<td>Negative scale [Mean (SD)]</td>
<td>16.1 (5.9)</td>
<td>12.5 (3.5)</td>
</tr>
<tr>
<td>General scale [Mean (SD)]</td>
<td>35.0 (6.6)</td>
<td>31.0 (6.2)</td>
</tr>
<tr>
<td>Total score [Mean (SD)]</td>
<td>67.0 (13.8)</td>
<td>59.4 (9.1)</td>
</tr>
<tr>
<td>Duration of voices, years [Mean (SD)]</td>
<td>16.7 (12.1)</td>
<td>14.5 (9.7)</td>
</tr>
<tr>
<td>Duration of commands, years [Mean (SD)]</td>
<td>16.2 (11.6)</td>
<td>14.3 (10.9)</td>
</tr>
<tr>
<td>Severity of worst ever CH [Mean (SD)]</td>
<td>3.4 (0.9)</td>
<td>3.9 (0.3)</td>
</tr>
<tr>
<td>Compliance with worst ever CH [Mean (SD)]</td>
<td>0.8 (0.8)</td>
<td>1.2 (0.9)</td>
</tr>
<tr>
<td>Degree of distress with recent CHs [Mean (SD)]</td>
<td>2.5 (1.2)</td>
<td>3.2 (0.8)</td>
</tr>
</tbody>
</table>

*a n's vary depending on available data.

*b 4-point severity scale based on legislation in the State of Victoria, Australia in which 1 = slightly dangerous/serious (e.g., shop steal); 2 = moderately dangerous/serious (e.g., burglary); 3 = very dangerous/serious (e.g., assault, serious self-harm) and 4 = extremely dangerous/serious (e.g., murder, suicide) (see Shawyer et al., 2008).

*c Based on 3-point compliance scale (0 = none, 1 = partial, 2 = full).

d Equal variances not assumed.
Therapeutic alliance

There were no significant differences between the two groups means on the quality of the therapeutic alliance, as measured by the HAq-II, at both sessions 4 [TORCH: 89.2 (SD: 8.3), Befriending: 91.1 (SD: 15.8); t(24.4) = -0.5, p = .64] and 10 [TORCH: 90.5 (SD: 13.0), Befriending: 93.0 (SD: 15.5); t(36) = -0.5, p = .59].

Overall trial quality

The CTAM score for this trial was rated at 88/100 by an independent reviewer.

Blindness

Breaches in blindness occurred on 22 occasions across the 97 assessments where blindness was relevant. By changing the rater, corrective action was able to be taken in all but 5 of these situations. One-tailed binomial tests, with a set at 0.05, and the probability of guessing allocation correct by chance set at 0.5, were used to assess the blindness of raters for the assessments that remained masked. A comparison of raters guesses of treatment (any) vs waitlist showed these were no better than chance, either immediately before, the first endpoint assessment, p = .56 (15/26 correct) or immediately after, p (exact) = 0.31 (15/24 correct). Following the cessation of waitlist allocations, raters guesses of treatment group (TORCH vs Befriending) for the remaining participants were also no better than chance immediately before the first endpoint assessment, p = .04 (3/10 correct), and immediately after, p (exact) = 1.00 (5/10 correct). Raters guesses of treatment group (TORCH vs Befriending) for waitlist participants at post therapy were no better than chance, before, p (exact) = 0.45 (6/16 correct) or after, p (exact) = 1.00 (8/15 correct) the post therapy assessment. At the final follow-up assessment, raters guesses at treatment group remained no better than chance immediately before, p = .10 (20/39 correct) and after the final follow-up assessment p = .74 (20/37 correct). Mean confidence ratings for raters guesses across assessment points ranged from 5 to 26%.

Outcomes

TORCH vs Befriending

Table 2 provides the means and between-group effect sizes associated with the primary and secondary outcome measures from baseline to endpoint and follow-up for TORCH and Befriending.

Primary hypothesis. Compliance and coping with command hallucinations: Compliance with harmful command hallucinations did not prove viable as an outcome measure. Although 93% (40/43) of the sample had experienced harmful commands in the past, only 64% (25/39) of this group had ever complied and, at baseline, less than half the sample (18/43) had complied to any degree with harmful commands in over the previous 4–6 months. Our primary measures were therefore limited to confidence to resist obeying harmful commands and confidence in coping with commands. As indicated in Table 2, no differences were found between or within the TORCH and Befriending groups on confidence to resist harmful commands at endpoint, F(1,35) = 0.05, p = .82, or follow up, F(1,30) = 0.02, p = .88. No differences were also found between TORCH and Befriending on confidence in coping with command hallucinations, at endpoint F(1,36) = 0.89, p = .35 or follow up, F(1,33) = 0.01, p = .91, however within-group analysis showed that both groups improved in confidence at endpoint relative to baseline. These improvements, however, were not sustained at follow up.

Secondary hypotheses. As shown in Table 2, none of the between-group results reached significance in this small study. However, analysis of within-group effects using paired t-tests showed improvements for both groups across different sets of variables. While TORCH participants showed significant improvements across the PANSS variables and the Modified GAF, significant improvement on the distress variables was observed only in the Befriending group. Both groups showed improvement in disruption (PSYRATS) and quality of life variables, although TORCH showed more robust improvement with respect to the latter. Improvements in

Table 2

<table>
<thead>
<tr>
<th>Measures</th>
<th>TORCH</th>
<th>Befriending</th>
<th>Cohen’s d&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Cohen’s d&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline (n = 18–20)</td>
<td>Endpoint (n = 18–20)</td>
<td>Follow up (n = 16–20)</td>
<td>Between group</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>endpoint</td>
</tr>
<tr>
<td>Primary confidence to resist</td>
<td>55.0 (29.8)</td>
<td>66.5 (24.7)</td>
<td>60.1 (29.9)</td>
<td>–.07</td>
</tr>
<tr>
<td>harmful CHs</td>
<td></td>
<td></td>
<td></td>
<td>–.05</td>
</tr>
<tr>
<td>Confidence in coping</td>
<td>49.3 (26.3)</td>
<td>65.0 (18.9)*</td>
<td>62.2 (22.6)</td>
<td>–.31</td>
</tr>
<tr>
<td>with CHs</td>
<td></td>
<td></td>
<td></td>
<td>.04</td>
</tr>
<tr>
<td>Secondary</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PANSS positive</td>
<td>16.3 (3.6)</td>
<td>14.5 (3.7)*</td>
<td>14.6 (3.4)</td>
<td>.19</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>15.7 (3.6)</td>
<td>.05</td>
</tr>
<tr>
<td>PANSS negative</td>
<td>16.3 (5.9)</td>
<td>14.8 (5.4)</td>
<td>14.0 (5.6)*</td>
<td>.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>12.1 (3.4)</td>
<td>-.02</td>
</tr>
<tr>
<td>PANSS general</td>
<td>35.3 (6.7)</td>
<td>30.9 (8.8)**</td>
<td>31.7 (7.2)*</td>
<td>.43</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>30.3 (6.0)</td>
<td>-.03</td>
</tr>
<tr>
<td>PANSS total</td>
<td>67.9 (13.6)</td>
<td>60.1 (15.9)**</td>
<td>60.2 (12.9)**</td>
<td>.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>58.1 (8.4)</td>
<td>.09</td>
</tr>
<tr>
<td>GAF</td>
<td>41.7 (12.2)</td>
<td>44.6 (11.3)</td>
<td>46.4 (10.7)*</td>
<td>.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>48.7 (9.5)</td>
<td>-.21</td>
</tr>
<tr>
<td>Distress (CHs)</td>
<td>2.4 (1.2)</td>
<td>2.2 (1.0)</td>
<td>2.3 (0.8)</td>
<td>-.40</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.3 (0.7)</td>
<td>.14</td>
</tr>
<tr>
<td>Distress (PSYRATS)</td>
<td>5.1 (2.5)</td>
<td>4.5 (2.5)</td>
<td>4.4 (2.5)</td>
<td>-.37</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4.9 (2.7)</td>
<td>-.06</td>
</tr>
<tr>
<td>Disruption (PSYRATS)</td>
<td>1.7 (0.8)</td>
<td>1.5 (0.8)</td>
<td>1.4 (0.7)*</td>
<td>-.52</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.7 (0.6)</td>
<td>-.06</td>
</tr>
<tr>
<td>QOL (feelings)</td>
<td>42.8 (11.3)</td>
<td>45.4 (8.1)</td>
<td>45.2 (11.8)</td>
<td>.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>45.1 (11.8)</td>
<td>-.48</td>
</tr>
<tr>
<td>QOL (general activities)</td>
<td>43.3 (10.6)</td>
<td>48.4 (9.3)*</td>
<td>46.2 (8.5)</td>
<td>.43</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>44.4 (13.0)</td>
<td>.14</td>
</tr>
<tr>
<td>Life satisfaction</td>
<td>2.7 (1.1)</td>
<td>3.3 (1.0)*</td>
<td>3.4 (1.0)**</td>
<td>.27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.9 (1.2)</td>
<td>.20</td>
</tr>
<tr>
<td>Life enjoyment</td>
<td>2.7 (0.9)</td>
<td>2.9 (0.7)</td>
<td>3.1 (0.8)*</td>
<td>.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.6 (1.0)</td>
<td>-.19</td>
</tr>
</tbody>
</table>

*p < .05; **p < .01; ***p < .001 based on paired t-tests compared with baseline.

<sup>a</sup> n varies depending on available data.
<sup>b</sup> Cohen’s d derived from ANCOVA statistics: marginal mean TORCH – marginal mean Befriending/square root of MSerror; positive values in direction of improvement for TORCH.
Befriending were mainly present at endpoint only while improvements in TORCH tended to emerge or remain at follow-up.

Subjective feedback. Based on data from the Client Satisfaction Questionnaire (n = 42), mean levels of satisfaction were similarly high across both treatment groups, t(31.9) = 1.6, p = .12 [M (SD); TORCH - 27/32 (2.9); Befriending - 25/32 (5.7)] though the feedback from Befriending participants showed significantly greater variability.

Mean ratings of problem improvement were significantly higher for TORCH participants compared to Befriending participants, (t(40) = 2.4, p = .02; M: 4.3/5 vs 3.8/5). Ninety percent of TORCH participants said their sessions made the problem of command hallucinations “better” or “much better”, compared to 59% of Befriending participants. With respect to emotional response to therapy, the majority of participants reported that their therapy sessions made them feel “better” or “much better” (85%) and there were no significant differences between groups on mean scores, (t(38) = .8, p = .40; M: 4.3/5 vs 4.1/5). Ninety-five percent of TORCH participants said their sessions made them feel “better” or “much better” compared to 76% of Befriending participants.

Process measures. Table 3 provides the means and between-group effect sizes associated with the process measures from baseline to endpoint and follow-up for TORCH and Befriending. Although there were no significant group effects, some different patterns of results were seen based on the within group analysis. Surprisingly, while the TORCH group showed improvement in their acceptance of general auditory hallucinations at follow-up, the Befriending group showed significant improvement in their acceptance of command hallucinations - with a moderate-large effect size, this group difference approached significance on ANCOVA (p = .07). TORCH participants also showed improvement in insight at endpoint; however, this was not maintained at follow up. Both groups showed significant reductions in BAVQ Omnipotence scores.

Waitlist vs treatment (TORCH and Befriending combined)

Clinical and demographic characteristics for the Treatment and Waitlist groups are provided in Table 4. There were no baseline differences in any of the variables.

The waitlist period (4 months) reflected the minimum expected therapy period, and was thus designed to provide only an approximation for the time therapy was expected to take (4–6 months). Not surprisingly then, the mean time between the pre and post therapy assessment of 7.1 (1.9) months significantly exceeded the time between the pre and post waitlist assessment of 5.3 (1.0) months, t(35.9) = 3.95, p = .00.

Table 5 provides the means and between-group effect sizes associated with the primary and secondary outcome measures from baseline to endpoint for Treatment and Waitlist. Examining between-groups comparisons, significant group differences favouring the Treatment group were found for a number of the measures including confidence in coping with command hallucinations, PANSS negative, PANSS general, PANSS total, Quality of Life (general activities) and life satisfaction. Further, the presence of medium effect sizes or above for the majority of variables strongly suggest differential effects favouring Treatment. Within the Treatment group, improvements were evident on paired t-tests on most of the measures while the Waitlist group either remained the same or deteriorated (PANSS negative and Quality of Life general activities).

Discussion

The aim of the study was to assess the efficacy of an acceptance-based CBT treatment for command hallucinations by comparison with an active comparison condition, Befriending. Participants engaged well with both treatments with equal levels of exposure to interventions and low attrition across the sample. Despite TORCH participants subjectively reporting greater improvement in command hallucinations compared to Befriending participants, the study found no significant group differences in primary and secondary outcome measures based on blinded assessment data. Within-group analyses, however, suggested that both treatments had positive effects with a differential pattern of outcomes observed across the two conditions. A comparison between the combined treatments and waitlist provided further indication that having either treatment was better than not having any treatment.

In comparing the results of this trial to the strongly positive outcomes reported by Trower et al. (2004) some substantial differences in design and population sampled should be noted. Trower et al. recruited participants from forensic and non-forensic services on the basis of having complied recently with severe commands. In terms of CBT trials for psychosis, their sample was unique and their effect size for positive symptoms is substantially higher than has been found in any other CBT for psychosis trial (Wykes et al., 2007). In contrast, participants in the trial reported here were exclusively from non-forensic mental health services, much less severely affected and possibly more typical of patients experiencing...
command hallucinations residing in the community. Our selection criteria were also broader: having experienced distressing or problematic command hallucinations in the previous 6 months, regardless of compliance. At baseline, only 42% of sample had complied to any degree with harmful command hallucinations in over the previous 4–6 months. Thus, while some participants’ main concern was compliance with harmful command hallucinations, others were disturbed by other elements of the experience such as the distress associated with ongoing resistance or frequent commands to perform everyday harmless acts. For a substantial proportion of participants at baseline (43%), other problems and issues were viewed as more important than command hallucinations. These more subtle and varied concerns resulted in a heterogeneous sample and allowed for less opportunity for clear and direct change related to command hallucinations to be observed across the group. Certainly, the primary outcome of resistance to harmful command hallucinations proved to be of less relevance for this group of stable outpatients than had been anticipated. This is not dissimilar to the experience reported in other psychosis trials where positive symptoms are not necessarily of most concern to participants and where the wide variation in concerns results in reduced power across different outcome measures (Farhall, Freeman, Shawyer, & Trauer, 2009; Peters et al., 2010).

The present trial also incorporated some important refinements including an active, well-validated control condition and tracking and verification of blindness. Added to this, we had low attrition.

Table 4
Demographic and clinical characteristics of Treatment and Waitlist groups at baseline.

<table>
<thead>
<tr>
<th>Demographic or Clinical Category</th>
<th>Treatment (n = 23–26)a</th>
<th>Waitlist (n = 12–17)</th>
<th>Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years [Mean (SD)]</td>
<td>39.9 (10.9)</td>
<td>39.6 (8.7)</td>
<td>t(41) = 0.11; NS</td>
</tr>
<tr>
<td>Sex (male/female)</td>
<td>13/13</td>
<td>11/6</td>
<td>χ2(1) = 0.90; NS</td>
</tr>
<tr>
<td>NART IQ [Mean (SD)]</td>
<td>100.0 (14.7)</td>
<td>98.9 (12.3)</td>
<td>t(38) = 0.24; NS</td>
</tr>
<tr>
<td>Years of education [Mean (SD)]</td>
<td>11.9 (2.2)</td>
<td>12.0 (2.6)</td>
<td>t(41) = −0.21; NS</td>
</tr>
<tr>
<td>DSM-IV Diagnosis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>18</td>
<td>13</td>
<td>Fisher’s exact test NS</td>
</tr>
<tr>
<td>Schizoaffective Disorder</td>
<td>7</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Mood disorder with psychotic features</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Duration of illness, years [Mean (SD)]</td>
<td>15.1 (10.8)</td>
<td>13.9 (7.4)</td>
<td>t(36) = 0.37; NS</td>
</tr>
<tr>
<td>Modified GAF score [Mean (SD)]</td>
<td>46.1 (10.6)</td>
<td>42.3 (15.3)</td>
<td>t(26.3)0.089; NS</td>
</tr>
<tr>
<td>Chlorpromazine equivalent dose (mg) [Mean (SD)]</td>
<td>774.9 (546.1)</td>
<td>705.4 (335.1)</td>
<td>t(41) = 0.47; NS</td>
</tr>
<tr>
<td>Involuntary status (yes/no)</td>
<td>4/22</td>
<td>0/16</td>
<td>Fisher’s exact test NS</td>
</tr>
<tr>
<td>PANSS score</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive scale [Mean (SD)]</td>
<td>16.1 (3.2)</td>
<td>16.2 (4.7)</td>
<td>t(41) = −0.13; NS</td>
</tr>
<tr>
<td>Negative scale [Mean (SD)]</td>
<td>13.3 (3.8)</td>
<td>13.4 (5.7)</td>
<td>t(41) = 0.10; NS</td>
</tr>
<tr>
<td>General scale [Mean (SD)]</td>
<td>31.8 (6.4)</td>
<td>33.3 (10.1)</td>
<td>t(41) = −0.61; NS</td>
</tr>
<tr>
<td>Total score [Mean (SD)]</td>
<td>61.1 (10.2)</td>
<td>63.0 (17.5)</td>
<td>t(23.2)0.040; NS</td>
</tr>
<tr>
<td>Duration of voices, years [Mean (SD)]</td>
<td>17.6 (12.3)</td>
<td>12.1 (7.1)</td>
<td>t(39.9)0.163; NS</td>
</tr>
<tr>
<td>Duration of commands, years [Mean (SD)]</td>
<td>16.7 (12.2)</td>
<td>12.4 (8.2)</td>
<td>t(23) = 1.09; NS</td>
</tr>
<tr>
<td>Severity of worst ever CH [Mean (SD)]</td>
<td>3.7 (0.6)</td>
<td>3.6 (0.9)</td>
<td>t(35) = 0.51; NS</td>
</tr>
<tr>
<td>Compliance with worst ever CH [Mean (SD)]</td>
<td>1.0 (0.9)</td>
<td>0.9 (0.8)</td>
<td>t(37) = 0.38; NS</td>
</tr>
<tr>
<td>Degree of distress with recent CHs [Mean (SD)]</td>
<td>3.1 (0.9)</td>
<td>2.6 (1.1)</td>
<td>t(41) = 0.179; NS</td>
</tr>
</tbody>
</table>

a n’s vary depending on available data.

b 4-point severity scale based on legislation in the State of Victoria, Australia in which 1 = slightly dangerous/serious (e.g., shop steal); 2 = moderately dangerous/serious (e.g., burglary); 3 = very dangerous/serious (e.g., assault, serious self-harm) and 4 = extremely dangerous/serious (e.g., murder, suicide) (see Shawyer et al., 2008).

c Based on 3-point compliance scale (0 = none, 1 = partial, 2 = full).

d Equal variances not assumed.

Table 5
Means (standard deviations) and effect sizes for outcome variables for Treatment and Waitlist from baseline to endpoint.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Treatment (n = 20–24)b</th>
<th>Waitlist (n = 19–24)</th>
<th>Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confidence to resist harmful CHs</td>
<td>54.3 (34.0)</td>
<td>69.1 (28.9)*</td>
<td></td>
</tr>
<tr>
<td>Confidence in coping with CHs</td>
<td>53.7 (26.4)</td>
<td>70.0 (22.9)**</td>
<td></td>
</tr>
<tr>
<td>PANSS positive</td>
<td>16.5 (2.9)</td>
<td>14.2 (2.9)**</td>
<td></td>
</tr>
<tr>
<td>PANSS negative</td>
<td>13.2 (3.9)</td>
<td>12.7 (3.3)</td>
<td></td>
</tr>
<tr>
<td>PANSS general</td>
<td>31.5 (6.5)</td>
<td>29.1 (6.7)</td>
<td></td>
</tr>
<tr>
<td>PANSS total</td>
<td>61.2 (10.1)</td>
<td>56.0 (10.2)*</td>
<td></td>
</tr>
<tr>
<td>GAF</td>
<td>47 (10.2)</td>
<td>46.6 (11.8)</td>
<td></td>
</tr>
<tr>
<td>Distress (Chs)</td>
<td>3.1 (0.9)</td>
<td>2.3 (1.1)</td>
<td></td>
</tr>
<tr>
<td>Distress (PSYRATS)</td>
<td>6.1 (1.6)</td>
<td>4.4 (2.9)**</td>
<td></td>
</tr>
<tr>
<td>Disruption (PSYRATS)</td>
<td>1.1 (0.5)</td>
<td>1.2 (0.8)**</td>
<td></td>
</tr>
<tr>
<td>QOL (subjective feelings)</td>
<td>41.6 (9.4)</td>
<td>46.2 (10.0)**</td>
<td></td>
</tr>
<tr>
<td>QOL (general activities)</td>
<td>44.3 (10.4)</td>
<td>47.0 (10.3)</td>
<td></td>
</tr>
<tr>
<td>Life satisfaction</td>
<td>2.7 (1.2)</td>
<td>3.3 (1.1)*</td>
<td></td>
</tr>
<tr>
<td>Life enjoyment</td>
<td>2.6 (0.8)</td>
<td>3.0 (0.9)*</td>
<td></td>
</tr>
</tbody>
</table>

a p < .05; **p < .01; ***p < .001 (within groups: results based on paired t-tests compared with baseline: between groups: results based on ANCOVA).

b Cohen’s d derived from ANCOVA statistics: marginal mean Treatment − marginal mean Waitlist/square root of MSerror; positive values in direction of improvement for Treatment.

c Significant deterioration.

Please cite this article in press as: Shawyer, F., et al., A randomised controlled trial of acceptance-based cognitive behavioural therapy for command hallucinations in psychotic disorders, Behaviour Research and Therapy (2012), doi:10.1016/j.brat.2011.11.007
and achieved statistically similar levels of exposure to the TORCH and Befriending treatments, which is important given that a number of previous trials have had less exposure to the control condition (Bendall et al., 2006; Haddock, Tarrier, et al., 1999; Lewis et al., 2002). The rigour of the trial is likely to have resulted in smaller between-group effect sizes. Building on an initial review by Tarrier and Wykes (2004), Wykes, Steel, Everitt, and Tarrier (2008) examined the relationship between CTAM score and effect size indicating that trials that employ better methodology tend to achieve smaller effect sizes. At 88, the CTAM score for the present trial was 1.5 standard deviation higher than the mean of 61.2 reported by Wykes et al. and also higher than the score of 71 reported by Wykes et al. for Trower et al. (2004).

The primary area of methodological weakness in the present study related to sample construction, which was less than the 30 per group indicated by our power analysis. The sample size was also less than would be desirable based on effect sizes reported in recent reviews (e.g., Lynch et al., 2010; Wykes et al., 2007) which are considerably lower than the review on which our power calculation was based (Rector & Beck, 2001) as well as others at the time (e.g., Gould, Mueser, Bolton, Mays, & Goff, 2001; Pilling et al., 2002). Despite recruitment efforts through 14 public mental health services, the rigour of the trial is likely to have resulted in a smaller between-group effect size compared with non-specific control interventions which is important given that sessions focusing attention away from problems and symptoms to topics that are less relevant is perhaps not surprising given that sessions focus attention away from problems and symptoms to topics that are positive or interesting for the participant. What is more surprising is the significant improvement in acceptance of command hallucinations by Befriending participants. Together with the significant findings for BAVQ omnipotence, these data suggest that the regular focus on real world activities and interests had direct effects in reducing the impact of commands on participants’ behaviour and attitudes. The non-significant findings in relation to acceptance of command hallucinations and distress in TORCH is disappointing, and may reflect the ‘being with’ nature of the therapist contact and expectancy in the context of CBT trials for psychosis. A number of recent reviews examining studies that compared CBT with non-specific treatments for psychosis have noted that the results from such trials are equivocal (Gaudiano, 2005, 2006; Jones, Cormac, Silveira da Mota Neto, & Campbell, 2004; Lynch et al., 2010; Tarrier & Wykes, 2004). In the most recent of these, Lynch et al. (2010) concluded that “CBT is no better than non-specific control interventions in the treatment of schizophrenia” (p. 9).

While Befriending has been viewed as merely a control for therapist contact and expectancy in the context of CBT trials for psychosis (Jackson et al., 2008; Sensky et al., 2000; Turkington & Kingdon, 2000), with any positive effects simply attributed to increased attention (Lynch et al., 2010), there have been recent suggestions that it is not a “no treatment” condition but an active and specific intervention in its own right. A recent examination of the content of Befriending sessions compared with CBT taken from an earlier trial conducted by Sensky et al. (2000) concluded that Befriending has a modus operandi that is different from that of CBT in terms of outcome — at least in the short term (Milne, Wharton, James, & Turkington, 2006). Milne found that Befriending was not a “diluted” form of CBT but that it appears to have beneficial effects through the provision of social support. An additional finding that symptom outcomes depend on symptom profile at baseline (Samarasekera et al., 2007) adds further support to the contention that this intervention involves more than just a generic effect of increased attention and expectation for improvement.

In the present trial, when the treatment groups were combined (n = 24) and compared with waitlist, significant group differences emerged for a range of factors including confidence in coping with commands, illness severity and quality of life. Further, in examining within group differences, the Treatment group improved on the majority of outcome measures while the Waitlist group showed no significant positive changes and even deteriorated on a couple of measures. Within-group analyses showed TORCH and Befriending groups both displaying improvements in confidence in coping with command hallucinations and life enjoyment together with reductions in disruption to life from auditory hallucinations and belief in the power of the voice(s) issuing commands.

In addition to these general improvements in both groups, the within group analyses related to the secondary and process measures suggest additional differential treatment effects. TORCH showed significant effects on a broader range of outcome measures including illness severity, global functioning and quality of life and process measures including acceptance of auditory hallucinations and insight. These effects tended to persist or emerge in the follow up period. It might be noted that the trend for Befriending across these measures was nearly always in the direction of improvement even if significance was not reached. Befriending alone showed significant effects on distress and these effects were only significant at endpoint. That Befriending showed significant short-term improvement on measures of distress is perhaps not surprising given that sessions focus attention away from problems and symptoms to topics that are positive or interesting for the participant. What is more surprising is the significant improvement in acceptance of command hallucinations by Befriending participants. Together with the significant findings for BAVQ omnipotence, these data suggest that the regular focus on real world activities and interests had direct effects in reducing the impact of commands on participants’ behaviour and attitudes. The non-significant findings in relation to acceptance of command hallucinations and distress in TORCH is disappointing, and may reflect 1) the significantly higher distress associated with command hallucinations experienced in the Befriending group at baseline; 2) the challenge of deliberately contacting difficult experiences rather than using strategies to avoid them or 3) the challenge of integrating two different therapeutic approaches, potentially diluting the effect of both by providing an overly broad and complex therapy for this population (Mueser, Drake, & Bond, 1997).

Notwithstanding this disappointment, the overall pattern of findings suggests that both treatments had positive benefits compared to waitlist, rather than both being ineffective, and that TORCH and Befriending work in different ways. Indeed, some highly specific improvements observed in Befriending in the absence of any direct intervention underpin the need for caution when linking even highly specific changes in outcome measures to the direct results of treatment when there is no active comparison group. Anecdotally, although some Befriending participants expressed a degree of frustration about not being able to discuss issues and problems, there were many positive comments from participants that endorse the value of Befriending in terms of social support and confidence and that reflect the “being with” as contrasted with the “done to” nature of this therapy (Samarasekera et al., 2007) (see Appendix).

Conclusion

With the quality of RCT trials rising substantially, it is becoming clear that CBT has more modest effects than the early studies...
promised and that non-specific treatment factors may be more important than was earlier thought (Mueser & Noordsy, 2005). It appears that much needs to be done to refine CBT-based treatments to improve on the effects of social support and unconditional positive regard. Although it is clear that acceptance-based CBT is well-tolerated and subjectively helpful to individuals with psychosis, there is limited evidence from this trial to indicate that integrating acceptance approaches with CBT improves outcomes over and above social support. However, given that the trial was underpowered, that the pattern of within group improvements varied across treatment condition, and that significant between-group changes emerged when the combined groups were compared to TAU, further research examining both acceptance-based CBT and Befriending with much larger sample sizes is warranted. In the context of the high costs of training associated with CBT for psychosis (Jones et al., 2004), and the importance of the social milieu in schizophrenia (Penn et al., 2004), it may be especially timely to investigate Befriending in more detail as a treatment for psychosis rather than a mere control for treatment (Jorm, 2005; Milne et al., 2006).

Declaration of interest

None.

Acknowledgements

This research was funded by grants from the National Health and Medical Research Council of Australia (Grant 251730) and the Rebecca L Cooper Medical Research Foundation. Therapist training in mindfulness was funded by grants from Novartis Pharmaceuticals Australia Pty Ltd and Eli Lilly Australia Pty Ltd. These funding sources had no involvement the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the paper for publication.

We would like to thank Sarah Bendall for providing initial instruction in Befriending and Nicholas Tarrier for rating the trial using the CTAM. We would also like to acknowledge the support of mental health services within the following organisations: Austin Health, Bendigo Health, Goulburn Valley Health, Peninsula Health, Melbourne Health, Mercy Health, Southern Health, and St. Vincent’s Health. Most of all, we would like to express our gratitude to our participants who made the trial possible.

Australian and New Zealand Clinical Trials Registry Number: 12605000440628

Appendix

Examples of comments from Befriending participants

“it gave me a more positive outlook of myself and the world around me”

“good ideas about life”

“scope to look at other things rather than voices”

“how important social interaction is in communicating your thoughts and feelings to someone”

“the realization that there is a lot of good things in my life to talk about”

“it helped to look at happier times”

“It was refreshing not to have to talk about upsetting things”

“I'm less isolative”

“the conversations make me think...make me communicate”

“I'm going to miss these conversations...I can have conversations with the Lord but they're a bit one-sided”

“meeting with you [therapist] has been really positive — it makes me think to myself people believe in me because you seem to believe in me whereas I've met people in the system...who don't seem to give a damn”

“you've helped — I feel I've finally got somewhere and I can get people to listen to me and I can talk to people whereas before I was sort of shut in — I was just sort of just sitting around with the voices pounding...I answered some questions too [in Bible study group] — its only about the first or second time I've actually spoken up...I think things will be better from now on”

Comments from Medical records

“Likes participating in Befriending activities. Reduction in symptoms — less auditory hallucinations currently which [participant] attributes to her increase in activities”

“Noted a decrease in periods of feeling sad with lowered mood. Explored contributing factors. [Participant] cited [Befriending]... reported that Befriending has helped build confidence and provided encouragement re social activities. Also noted he is less isolative”

References


