

Cognitive therapy for command hallucinations: randomised controlled trial

PETER TROWER, MAX BIRCHWOOD, ALAN MEADEN, SARAH BYRNE, ANGELA NELSON and KERRY ROSS

Background Command hallucinations are a distressing and high-risk group of symptoms that have long been recognised but little understood, with few effective treatments. In line with our recent research, we propose that the development of an effective cognitive therapy for command hallucinations (CTCH) would be enhanced by applying insights from social rank theory.

Aims We tested the efficacy of CTCH in reducing beliefs about the power of voices and thereby compliance, in a single-blind, randomised controlled trial.

Method A total of 38 patients with command hallucinations, with which they had recently complied with serious consequences, were allocated randomly to CTCH or treatment as usual and followed up at 6 months and 12 months.

Results Large and significant reductions in compliance behaviour were obtained favouring the cognitive therapy group (effect size=1.1). Improvements were also observed in the CTCH but not the control group in degree of conviction in the power and superiority of the voices and the need to comply, and in levels of distress and depression. No change in voice topography (frequency, loudness, content) was observed. The differences were maintained at 12 months' follow-up.

Conclusions The results support the efficacy of cognitive therapy for CTCH.

Declaration of interest None.

Command hallucinations are high-risk, distressing and relatively common symptoms of schizophrenia (Beck-Sander *et al*, 1997; Shawyer *et al*, 2003). Shawyer *et al* (2003) find a median prevalence of 53% and median prevalence of compliance of 31%. There are few treatment approaches and none tested systematically. Indications are that command hallucinations feature strongly in those considered 'treatment resistant', and even hospitalisation is not necessarily a barrier to compliance (e.g. Jones *et al*, 1992). However, progress has been made in the development of cognitive therapy for hallucinations in general, and we believe command hallucinations are particularly appropriate for a cognitive approach. In addition, social rank theory (Gilbert, 1992) can account for the cognitive content of the specific beliefs of command hallucination hearers (Byrne *et al*, 2003). Following the principles of social rank theory, the authors have developed cognitive therapy for command hallucinations (CTCH), which does not depend on reducing the experience of voices but on reducing the perceived power of voices to harm the individual and to motivate compliance (Birchwood *et al*, 2000). In the present paper we describe a single-blind, intention-to-treat randomised controlled trial in which we compare the efficacy of CTCH plus treatment as usual (TAU) with TAU alone, in a sample of participants with command hallucinations considered at high risk of further compliance by virtue of serious recent compliance. The main hypothesis (and primary outcome) was that by challenging key beliefs about the power of commanding voices, the CTCH group would show a lower level of compliance and appeasement behaviour and an increase in resistance compared with the control group. The secondary outcomes were a lower conviction in the power and social rank superiority of voices and the need to comply, and a reduction in distress and depression. No change in the severity

of positive symptoms or the topography of voices (frequency, loudness, content) was predicted.

METHOD

Recruitment and procedure

The participants were recruited from local mental health services in Birmingham and Solihull, Sandwell and a West Midlands semi-secure unit for offenders with mental illness. Inclusion criteria were that patients conformed to an ICD-10 diagnosis of schizophrenia or related disorder with command hallucinations for at least 6 months (World Health Organization, 1992). Participants were required to have a recent history of compliance with, and appeasement of, voices with 'severe' commands, including harm to self, others or major social transgressions. Patients were excluded if they had a primary organic or addictive disorder.

All aspects of recruitment, screening and outcome assessment were organised and administered by an experienced research associate (A.N.) between September 2000 and July 2002. All patients referred were offered an interview to establish eligibility and to obtain consent, and a further interview for eligible patients for assessment with the outcome measures (see below). Eligible and consenting patients were then randomly assigned to CTCH or TAU by means of a computerised random number generator administered by the Birmingham Clinical Trials Unit independent of the research team, to ensure the research associate was blind to the allocation at baseline and post-testing. Participants were post-tested at 6 months after CTCH or TAU and again at 12 months' follow-up.

Power calculations were based on previous cognitive therapy trials which suggest a relatively large effect size of 0.65 post treatment and 0.93 at 12 months' follow-up (Gould *et al*, 2001; Cormac *et al*, 2002); predicting a reduction from 95% with at least partial compliance to 50% would require a sample size of 23 in each of two groups to achieve a power of 0.9 with alpha=0.05.

Measures

Measures of cognitions and behaviours, symptoms and affect relevant to the hypotheses were given at pre-test, post-test and follow-up. The cognitive and behavioural measures were as follows.

Cognitive Assessment Schedule

The Cognitive Assessment Schedule (CAS; Chadwick & Birchwood, 1995) is a measure of the individual's feelings and behaviour in relation to the voice, and beliefs about the voice's identity, power, purpose or meaning and the likely consequences of obedience or resistance.

Beliefs About Voices Questionnaire

The Beliefs About Voices Questionnaire (BAVQ; Chadwick & Birchwood, 1995; Chadwick *et al.*, 2000) measures key beliefs about auditory hallucinations, including benevolence, malevolence and two dimensions of relationship with the voice: 'engagement' and 'resistance'.

Voice Compliance Scale

The Voice Compliance Scale (VCS; Beck-Sander *et al.*, 1997) is an observer-rated scale to measure the frequency of command hallucinations and level of compliance/resistance with each identified command. The VCS was completed in two stages. First, the trial assessor (A.N.) used a structured interview format to obtain from each client a description of all those commands and associated behaviours (compliance or resistance) within the previous 8 weeks where they felt compelled to respond. The assessor then interviewed either a key-worker or relative to corroborate the information, and where there was a discrepancy, recorded the worst behaviour mentioned by either party. To further corroborate the accuracy of the information, and to ensure blindness to the allocation, a behavioural scientist (K.R.) was employed 6 months post-trial to check the record of commands and associated compliance and resistance behaviours obtained from interview against the case notes. Concordance was 100% for severe commands, giving confidence in the reliability of the data. Second, the assessor then classified each behaviour as: (1) neither appeasement nor compliance; (2) symbolic appeasement, i.e. compliant with innocuous and/or harmless commands; (3) appeasement, i.e. preparatory acts or gestures; (4) partial compliance with at least one severe command; (5) full compliance with at least one severe command. The behaviours were also independently and blindly rated using the information collated from the informants by three of the authors (M.B., A.M. and P.T.), and interrater reliability (Fleiss,

1981) for three judges using the whole sample at 6 months was found to be good ($\kappa=0.78$). Discrepancies were resolved by discussion and taking the mean rating. The scale also has good construct validity (see Results).

Voice Power Differential scale

The Voice Power Differential scale (VPD; Birchwood *et al.*, 2000) measures the perceived relative power differential between voice and voice hearer, with regard to the components of power, including strength, confidence, respect, ability to inflict harm, superiority and knowledge. Each is rated on a five-point scale and yields a total power score.

Omniscience Scale

The Omniscience Scale (Birchwood *et al.*, 2000) measures the voice hearer's beliefs about the knowledge of their voice regarding personal information.

Other rating scales

Measures for symptoms and distress include:

- (a) Positive and Negative Syndrome Scale (PANSS; Kay *et al.*, 1987). This is a widely used, well established and comprehensive symptom rating scale measuring mental state.
- (b) Psychotic Symptom Rating Scales (PSYRATS; Haddock *et al.*, 1999). This measures the severity of a number of dimensions of auditory hallucinations and delusions, including the amount and intensity of distress associated with these symptoms.
- (c) Calgary Depression Scale for Schizophrenia (CDSS; Addington *et al.*, 1993). This is designed specifically for assessment of the level of depression in people with a diagnosis of schizophrenia.

All the above measures have satisfactory psychometric properties, reported in the journal articles cited.

Treatment groups

Consenting participants were assigned randomly to either TAU or CTCH plus TAU for a period of 6 months. The research associate responsible for outcome evaluation was blind to group allocation (A.N.) and participants were instructed not to disclose their allocation.

TAU

This was delivered by community mental health teams. A detailed breakdown of the services received by the control and treatment groups during the trial and 1 year before the trial are shown in Table 1. This shows that TAU was extensive, involving 18 categories of service and admissions.

Medication was recorded 12 months before, and during, the trial.

CTCH

The key foci of the assessment, formulation and intervention are four core dysfunctional beliefs (and their functional relation to behaviour and emotion) that define the client-voice (social rank) power relationship: that the voice has absolute power and control; that the client must comply or appease, or be severely punished; the identity of the voice (e.g. the Devil); and the meaning attached to the voice experience (e.g. the client is being punished for past bad behaviour). Using the methods of collaborative empiricism and Socratic dialogue, the therapist seeks to engage the client to question, challenge and undermine the power beliefs, then to use behavioural tests to help the client gain disconfirming evidence against the beliefs. These strategies are also used to build clients' alternative beliefs in their own power and status, and finally, where appropriate, to explore the origins of the schema so clients have an explanation for why they developed those beliefs about the voice in the first place. These interventions are designed to enable the individual to break free of the need to comply or appease, and thereby reduce distress. The CTCH was given in line with the protocol developed by M.B. and P.T. by one of the authors (S.B.) – a clinical psychologist experienced in cognitive therapy and supervised in CTCH by A.M., M.B. and P.T. A behavioural scientist (K.R.) independent of the trial rated a random selection of early, middle and late audiotaped sessions (13 in total) using the Cognitive Therapy Checklist (Haddock *et al.*, 2001). The mean rating was 54 (range 51–56), indicating a very high level of concordance. All sessions achieved 4 or more on each subsection on a scale where 0=inadequate, 6=excellent and 4=good. The scale is divided into general (agenda, feedback, understanding, interpersonal effectiveness and collaboration) and specific (guided discovery, focus on key cognitions, choice of intervention, homework

Table 1 Service consumption before and during the trial: proportion of patients using services, categorised by treatment group

	Year before trial		During trial	
	TAU	CTCH	TAU	CTCH
	%	%	%	%
Services				
Out-patients	85	89	100	89
CPN	50	33	60	56
Day centre	45	38	35	28
Social worker	0	22	25	11
Supported accommodation	30	22	55	28
Support worker	25	50	45	11
Community drug team	5	6	5	0
Probation officer	5	6	5	6
Occupational therapist	35	22	20	11
Psychologist	15 ¹	16 ¹	5 ²	6 ²
Respite care	0	6	5	6
Home treatment team	22	6	20	6
Art therapy	10	6	5	0
Voices group	0	6	0	0
ECT	0	0	0	6
Admissions				
Informal	20	22	15	22
Section 2	10	0	5	6
Section 3	15	0	15	0
Guardianship Order	0	6	0	6

ECT, electroconvulsive therapy; CPN, community psychiatric nurse; CTCH, cognitive therapy for command hallucinations; TAU, treatment as usual.

1. Psychological input in the treatment group included anger management, childhood trauma and symptom management. Psychological input in the control group constituted anxiety and anger management.

2. Psychological input for the treatment group constituted anger management. Psychological input in the control group constituted anxiety management.

and quality of intervention). The treatment protocol is described fully in Byrne *et al* (2003).

Neuroleptic medication

The daily dose of neuroleptic medication at baseline, 6 and 12 months was recorded from case notes and converted to chlorpromazine (CPZ) equivalents using the conversion described in the *British National Formulary* (British Medical Association & Royal Pharmaceutical Society of Great Britain, 2003). Conversion from atypical to typical (CPZ) medication is to a degree arbitrary, but we employed the same formula for both groups; thus statistical comparison between groups would be unaffected.

Statistical analysis

Hypotheses were tested using the Generalized Linear Interactive Modelling Program

(GLIM) in the Statistical Package for the Social Science for Windows, version 10. The statistical model was treatment group \times time, with repeated measures on the time factor. The test of each hypothesis focused on the interaction term. It is also of interest to determine if there are general trends across time in both groups (e.g. reduction in compliance) tested using the 'time' main-effect term of the GLIM analysis.

To test whether the intervention was effective, baseline *v.* 6 months measures were used; for maintenance of any treatment effects, baseline *v.* 12 months measures were used. Exact probability values were calculated.

RESULTS

Description of the sample

A total of 224 referrals were screened, from which 69 patients were identified as being

eligible for the study and were invited to participate. Of these, 31 refused consent, leaving a sample of 38 consenting to randomisation (Fig. 1). The sample included 24 men and 14 women, with a mean age of 35.5 years (s.d. 10.4). The sample was drawn from a broad ethnic base, including 27 (71%) White, 6 (16%) Black Caribbean and 4 (14%) other/South Asian patients. The clinical and demographic characteristics of the treatment and control groups are shown in Table 2. Those refusing consent did not differ from the participants on available data (gender, age, duration of illness).

Prescribed neuroleptic medication converted to CPZ equivalents is presented in Table 3. No difference was observed at baseline between the two groups ($F=2.0$, NS). At baseline, 13/18 (72%) in CTCH were prescribed atypicals, including 5 patients taking clozapine; in TAU, 13/20 were prescribed atypicals (65%), including 7 patients taking clozapine. A group \times time repeated measure analysis of variance conducted on the drug data confirmed no overall difference between groups, but found a group \times time interaction ($F=6.3$, $P=0.005$). Table 3 shows that this was due to a steady rise in prescribed neuroleptic drugs in the TAU group ($t=3.0$, $P<0.01$) and a small but significant decrease in the CTCH group ($t=2.3$, $P<0.05$). The numbers of participants receiving atypicals in either group were unchanged at follow-up.

Types of commands, compliance and forensic history

All patients reported two or more commands from the 'dominant' voice, at least one of which was a 'severe' command. The most severe commands were to kill self (25), kill others (13), harm self (12) and harm others (14). Less severe commands involved innocuous, everyday behaviour (wash dishes, masturbate, take a bath) and minor social transgressions (e.g. break windows, shout out loud, swear in public). Further details, including incidence and examples of compliance and appeasement of such commands for the sample as a whole, are shown in Table 4.

Participants were considered at high risk of compliance because 30 (79%) had complied, 14 (37%) had appeased, and 29 (76%) had expressed the fear that the voices would either harm or kill them or a family member if they did not comply.

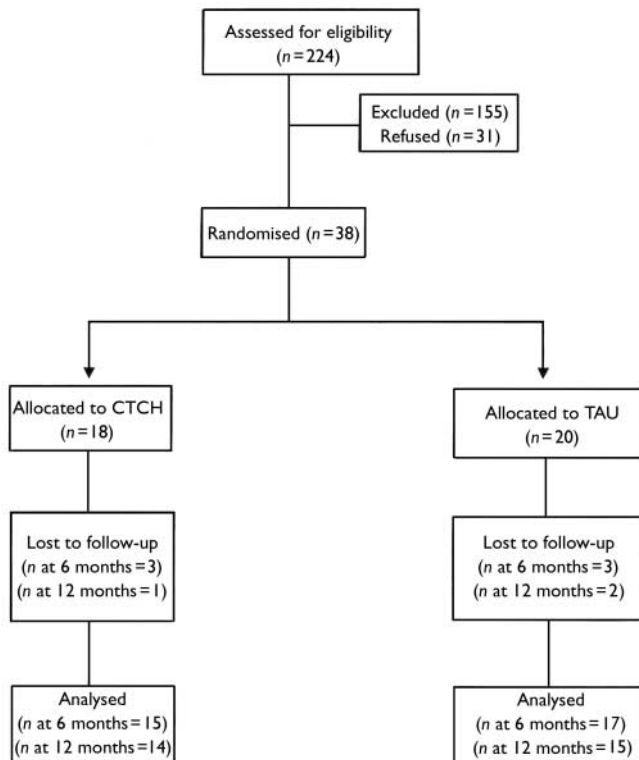


Fig. 1 CONSORT diagram.

The compliance rate is at the high end of the range for recent studies (Sawyer *et al*, 2003), because our sampling strategy involved identifying those considered to have recently complied.

Five participants in the sample had been prosecuted or cautioned for behaviour linked to voices' commands. This included causing actual bodily harm to a minor, grievous bodily harm, theft and common assault. Three participants had been hospitalised (two detained under the Mental Health Act, 1983), for attempting to kill someone in response to voices within the last 3 years.

A further indication of the severity of need in this sample was the heavy and prolonged consumption of TAU, both during the trial and as sampled 1 year before the trial. TAU involved 17 categories of service and admissions, as shown in Table 1.

Another indication of severity was the fact that at the time of consent to enter the trial, eight patients were hospitalised; two admissions were under Section 3 and one under Section 2 of the Mental Health Act (1983), and another five were informal admissions.

Allocation and flow of participants

As shown in the CONSORT diagram (Fig. 1), 38 of 69 eligible participants were

randomly allocated, 18 to CTCH and 20 to TAU. The CTCH and TAU groups did not significantly differ on any demographic, illness history or voice characteristics at baseline (see Table 2). The treatment group completed a median of 16 sessions. Five participants (27%) in the treatment group dropped out prematurely, attending between 4 and 12 sessions. This drop-out rate is comparable with other trials of this type (Norman & Townsend, 1999; Durham *et al*, 2003). The intention was to include all 18 CTCH participants at follow-up, but at 6 months three participants were lost to follow-up through withdrawal of consent, and a further one at 12 months. In the control group, two were lost to follow-up at 6 months (both died; one death was due to natural causes and the other to suicide) and two were lost at 12 months. There was no difference between groups in number lost to follow-up.

The impact of CTCH

Compliance with commands

The CTCH and TAU groups did not differ in compliance with commands at baseline, as measured by the VCS. There was a general effect of time, with both groups showing a reduction in compliance ($F=89.3$, $P<0.0001$); however, this was

Table 2 Clinical and demographic characteristics of the treatment and control groups

	CTCH	TAU
Age, years		
Mean (s.d.)	36.6 (10.3)	35.1 (10.4)
Range	17–56	19–60
Gender, n		
Male	10	14
Female	8	6
Ethnicity, n		
White	13	14
Black	4	3
Asian	1	1
Other	0	2
Diagnosis, n		
Schizoaffective	4	0
Schizophrenia	6	11
Paranoid schizophrenia	5	5
Personality disorder	1	2
Psychotic depression	1	2
OCD	1	0
PANSS score		
Positive scale		
Mean (s.d.)	21.9 (3.1)	20.8 (3.2)
Range	16–28	16–28
Negative scale		
Mean (s.d.)	20.8 (6.4)	21.5 (6.4)
Range	12–34	13–34
General psychopathology		
Mean (s.d.)	36.3 (6.6)	35.9 (6.7)
Range	26–47	28–49
Duration of voices, years		
Mean (s.d.)	13.4 (9.9)	10 (5.7)
Duration of commands, years		
Mean (s.d.)	8.8 (7.9)	8.6 (5.9)

CTCH, cognitive therapy for command hallucinations; OCD, obsessive-compulsive disorder; PANSS, Positive and Negative Syndrome Scale; TAU, treatment as usual.

Table 3 Changes in prescribed antipsychotic medication

	CTCH (n=17)	TAU (n=16)
	Mean (s.d.)	Mean (s.d.)
Chlorpromazine equivalent, mg/day		
Baseline	1181 (948)	779 (646)
6 months	1167 (958)	893 (601)
12 months	1127 (969)	1038 (657)

CTCH, cognitive therapy for command hallucinations; TAU, treatment as usual.

Table 4 Prevalence of and types of commands, compliance and appeasement in the whole sample

Command and prevalence	Example	Compliance	Appeasement
Command to kill self (treatment $n=12$, control $n=13$)	'Stab yourself' 'Slash your wrists' 'Overdose' 'Hang yourself' 'Gas yourself'	Nine patients had previously attempted suicide. One patient committed suicide during the trial	Seven patients used appeasement behaviours including holding a knife to their wrist, taking razor blades into the bath, collecting tablets and planning and executing suicide in their imagination
Command to kill others (treatment $n=6$; control $n=7$)	'Cut her throat' 'Go and kill someone' 'Kill the therapist' 'Kill your husband and daughter'	Four patients in the sample had attempted to kill someone, either by suffocation, poisoning or physical assault with a hammer	Three patients used appeasement behaviours including arming themselves with knives, baseball bats and an axe and making guns out of tinfoil
Command to harm self (treatment $n=9$; control $n=3$)	'Burn yourself' 'Cut yourself' 'Set yourself alight' 'Pour hot water on yourself' 'Go into the road'	Nine patients had harmed themselves in response to commands. This included cutting, swallowing nail polish remover or bleach, jumping in front of cars, walking on glass and setting oneself alight	Three patients used appeasement behaviours including picking at previous wounds, and standing on the kerb
Command to harm others (treatment $n=8$; control $n=6$)	'Touch your children' 'Kick them' 'Hit them' 'Beat that person up' 'Rape your neighbour'	Seven patients had harmed others in response to commands. This included hitting children, knocking them into furniture, scolding them, hitting people, attacking someone with a knife	Two patients used appeasement behaviours including hitting others with minimal force, and covert appeasement by thinking 'I'll do that later'

marked in the CTCH group ($F_{\text{interaction}}=4.8$, $P=0.036$; Table 5). These treatment gains were maintained at 12 month follow-up ($F_{\text{interaction}}=7.8$, $P<0.001$). To gauge the comparative size of this effect, at 6 months 39% of the TAU group scored 4 or 5 on the compliance scale (partial or full compliance) compared with 14% of the CTCH group. This compares with a baseline figure of 100% for the CTCH and 94% for the TAU groups. The effect size of CTCH was 1.1.

These findings were unaffected when change in neuroleptic dose over time was used as covariate. However, within the TAU group, the rise in neuroleptic dose was correlated ($r=0.46$, NS) with reducing compliance; and in the CTCH group, the reduction in medication use was correlated ($r=0.63$, $P<0.01$) with reducing compliance. At baseline and follow-up, there was no correlation between medication dose and compliance or voice power.

Beliefs about voices

Findings on the four main voice belief variables were as follows:

- Power (VPD). The CTCH group reported a large and significant reduction in the power of the dominant voice, compared with the TAU group, which showed no change ($F_{\text{interaction}}=19.4$, $P<0.0001$; Table 6). This effect of CTCH was maintained at 12 months follow-up ($F_{\text{interaction}}=15.1$, $P<0.001$).
- Malevolence (BAVQ). There was no impact of CTCH on the perceived malevolence of voices at 6 months ($F_{\text{interaction}}<1$, NS) or 12 months ($F_{\text{interaction}}<1$, NS).
- Omniscience (BAVQ). The belief in voices' omniscience declined significantly in the CTCH group but not in the TAU group ($F_{\text{interaction}}=3.9$, $P=0.05$). This pattern was maintained at 12 months ($F_{\text{interaction}}=6.3$, $P=0.02$).
- Perceived control (PSYRATS). Patients receiving CTCH showed a significant improvement in perceived control over voices, compared with the TAU group, which showed no change ($F_{\text{interaction}}=11.3$, $P=0.002$). This pattern was maintained at 12 months ($F_{\text{interaction}}=7.2$, $P=0.01$).

Changes in compliance and the perceived power of voices

Our theory predicts that it is the focus on the power of the voice that is responsible for the observed reductions in compliance behaviour. This was tested by comparing the two groups on compliance at follow-up and entering voice power (VPD) as covariate, predicting that the covariate will render the treatment effect non-significant. At 6 months, the two groups were significantly different in compliance (CVS: $F=7.3$, $P=0.011$); however, when power was entered as covariate, the difference was no longer significant ($F=2.6$, NS). At 12 months, the groups were significantly

Table 5 Mean scores (s.d.) on the Voice Compliance Scale showing impact of CTCH on compliance with commands

	CTCH	TAU
Baseline	4.7 (0.5)	4.8 (0.5)
Post-treatment	1.8 (1.2)	3.1 (1.4)
12 months	1.7 (1.1)	3.4 (1.6)

CTCH, cognitive therapy for command hallucinations; TAU, treatment as usual.

Table 6 Mean scores (s.d.) showing impact of CTCH compared to TAU on measures of voice beliefs, topography and distress

	CTCH			TAU		
	Baseline	6 months	12 months	Baseline	6 months	12 months
Voice power (VPD)	28.0 (4.7)	17.4 (6.7)	18.0 (6.5)	26.0 (7.8)	26.7 (5.3)	29.1 (6.2)
Malevolence (BAVQ)	23.2 (6.9)	18.3 (7.5)	21.7 (8.6)	19.6 (8.8)	19.8 (8.9)	20.7 (8.6)
Omniscience (BACQ)	11.4 (2.8)	9.1 (4.4)	9.1 (2.1)	9.8 (3.1)	10.2 (4.1)	10.0 (4.6)
Distress (PSYRATS)	3.4 (0.5)	2.2 (0.9)	2.7 (0.8)	3.2 (0.73)	2.9 (1.0)	3.0 (1.1)
Frequency (PSYRATS)	3.5 (0.5)	2.5 (1.0)	2.7 (1.3)	2.8 (1.1)	2.7 (1.1)	2.9 (1.0)
Loudness (PSYRATS)	2.6 (1.0)	2.0 (0.9)	2.6 (1.3)	2.2 (1.2)	2.4 (0.9)	2.3 (0.8)
Negative content (PSYRATS)	3.4 (0.6)	3.1 (1.1)	3.6 (0.7)	3.6 (0.5)	3.4 (0.8)	3.7 (0.6)
Control (PSYRATS)	3.7 (0.5)	2.7 (1.4)	2.6 (1.5)	3.4 (0.8)	3.6 (0.9)	3.6 (0.8)
Depression (CDSS)	9.7 (5.9)	8.0 (6.3)	8.1 (7.4)	8.9 (6.8)	9.3 (6.9)	12.6 (6.7)

CTCH, cognitive therapy for command hallucinations; TAU, treatment as usual; VPD, Voice Power Differential Scale; BAVQ, Beliefs About Voices Questionnaire; PSYRATS, Psychotic Symptom Rating Scale; CDSS, Calgary Depression Scale for Schizophrenia.

Table 7 Correlations between voice compliance, distress, power and omniscience of disobedience

	Distress	Power	Omniscience
Voice compliance	0.38*	0.46**	0.32
Distress (PSYRATS)		0.55**	0.47**
Power differential (VPD)			0.37

PSYRATS, Psychotic Symptom Rating Scale; VPD, Voice Power Differential Scale.

* $P < 0.05$; ** $P < 0.01$.

different ($F=9.8$, $P=0.004$), but again this disappeared when controlling for VPD at 12 months ($F < 1$, NS).

Distress and depression

Findings on the two key effect variables were as follows:

(a) Distress (PSYRATS). Intensity of distress fell significantly in the CTCH group at 6 months but not in the control group ($F_{\text{interaction}}=5.3$, $P=0.03$). By 12 months distress levels in the groups were no longer different ($F=2.7$, NS) but there was an overall lessening of distress over this period ($F=4.2$, $P=0.05$).

(b) Depression (CDSS). There was no change in depression scores with time ($F < 1$) and no interaction with treatment group ($F_{\text{interaction}}=1.3$, NS). However, by 12 months, depression had risen significantly in the TAU but not in the CTCH group ($F_{\text{interaction}}=7.3$, $P=0.012$). The baseline score of the whole sample (s.d. 6.4) indicates moderate depression (Addington *et al*, 1993).

Voice topography

Findings here were largely in line with predictions.

(a) Voice frequency (PSYRATS). Perceived voice frequency fell in the CTCH group compared with the TAU group ($F_{\text{interaction}}=6.8$, $P=0.022$), which did not change from baseline. This difference was not maintained at 12 months ($F_{\text{interaction}}=3.4$, NS).

(c) Voice content (PSYRATS). The reported negative content of voices did not change in either group with time (all $F < 1$).

Psychotic symptoms

Although change in psychotic symptoms was not predicted, a significant drop occurred in PANSS positive symptoms amounting to 3.7 points in the CTCH group, from a baseline of 21.8, and a small increase occurred in the control group ($F_{\text{interaction}}=12.6$, $P < 0.001$). Similarly, there was a small but consistent reduction in negative symptoms ($F_{\text{interaction}}=14.8$, $P=0.001$) and general psychopathology ($F_{\text{interaction}}=18.8$, $P < 0.001$) in the CTCH group (Table 7). These effects were

maintained at 12 months for positive symptoms ($F_{\text{interaction}}=14.2$, $P=0.001$), negative symptoms ($F_{\text{interaction}}=12.3$, $P=0.002$) and general psychopathology ($F_{\text{interaction}}=15.5$, $P=0.001$).

Within the PANSS positive scale, hallucinations showed a non-significant reduction at 6 months ($F=3.8$, $P=0.06$); however, by 12 months, no difference between the groups was observed ($F_{\text{interaction}}=1.46$, NS). In contrast, for the delusions sub-scale there was a reduction in the CTCH group at 6 months ($F_{\text{interaction}}=5.6$, $P=0.0025$), sustained at 12 months ($F_{\text{interaction}}=3.98$, $P=0.005$).

Within the general psychopathology scale, there were significant changes in anxiety at 6 months ($F_{\text{interaction}}=10.6$, $P=0.004$) and 12 months ($F_{\text{interaction}}=9.9$, $P=0.004$); in tension at 6 months ($F_{\text{interaction}}=5.1$, $P=0.03$); and in guilty thinking at 6 months ($F_{\text{interaction}}=4.6$, $P=0.042$).

Within the negative symptoms scale, there was a significant reduction in attention/concentration at 6 months ($F_{\text{interaction}}=13.2$, $P=0.001$), and disturbance of volition at 6 months ($F_{\text{interaction}}=6.2$, $P=0.019$) and 12 months ($F_{\text{interaction}}=15.5$, $P=0.001$).

There was no correlation between neuroleptic dose and PANSS positive symptoms at any point.

Construct validity of the VCS

Social rank theory applied to the experience of voices argues that compliance with a powerful dominant (voice) will vary as a function of: the power differential between the dominant (voice)

and subordinate (voice hearer); the distress or fear experienced; and the beliefs about non-compliance (see Gilbert, 1992).

We can put this critical aspect of our theory to the test. This will, in addition, serve to test the validity of the VCS if it correlates lawfully with these self-report scales.

The correlation matrix at 6 months (when the VCS has more variability) is shown in Table 7. Voice compliance was correlated significantly with both greater distress and power, with a trend for omniscience (multiple $R=0.55$, $P<0.01$).

DISCUSSION

It is important to reiterate that the people in this study were selected as being at 'high risk': they had complied with 'serious' commands to self-harm, harm others or to commit major social transgressions; they were highly distressed; and many were 'appeasing' the dominant voice in order to 'buy time' to avoid what they believed to be catastrophic consequences. Many had a history of forensic involvement, and all were supported by community teams who referred the patients because of perceived risk where clinicians acknowledged equipoise in their management.

Reduction in compliance

The data presented here suggest that CTCH, in the context of good quality and a high level of TAU services, exerts a major influence on the risk of compliance, reduces distress and prevents the escalation of depression, compared with TAU alone. Depression is known to be high in this group from previous research (Birchwood *et al*, 2000), confirmed in this study. Because of the selection criterion of recent compliance, it was likely that compliance behaviour would reduce over the 6-month and 12-month periods ('regression to the mean'); however, given the high risk status of this group, we could expect an increasing number of people complying with commands as further time elapses. Nevertheless, the 12-month clinical impact of CTCH was significant. Perhaps more importantly, the risk factors for compliance in the CTCH group had reduced markedly, particularly the perceived power of the voice, its omniscience and controllability, and the need to appease it (14% of the

CTCH group were appeasing or complying *v.* 53% of the TAU group).

Change in beliefs

In line with our prediction, neither the topography nor the negative content of voices shifted (according to the self-reports of participants on the PSYRATS), with the exception of a temporary reduction in perceived frequency during the first 6 months. This underlines our view (Birchwood & Spencer, 2002) that cognitive-behavioural therapy (CBT) is most effective with beliefs (delusional or otherwise), rather than the primary psychotic experience, in this case auditory hallucinations. The focus of CTCH is to change fundamentally the nature of patients' relationships with their voices by challenging the power and omnipotence of the voices, thus reducing the motivation to comply. However, if these treatment gains are sustained, the reduction of distress might well exert a beneficial influence on the frequency of voices. In a similar vein, we previously observed the similarity between the nature and content of voices and negative thoughts in depression (Gilbert *et al*, 2001); relieving depression in this sample could act to reduce the frequency and negative content of voices, although in the time scale observed here, only limited change was noted.

The reduction in PANSS positive symptoms was modest but consistent in the CTCH group, leading to a highly significant and sustained effect. The hallucinations sub-scale showed a non-significant decline over 6 months which disappeared at 12 months, once more underlining our contention that voice activity *per se* is not affected. The delusions sub-scale, in contrast, showed a significant reduction at 6 and 12 months. This could well reflect the observed changes in the perceived power of the persecutor, in this instance the voice. The PANSS general psychopathology score showed the largest and most sustained reduction, particularly social avoidance, attention and concentration.

Internal and external validity

Our primary dependent measure – compliance with commands – is not a straightforward concept (Beck-Sander *et al*, 1997), as compliance can include both covert as well as overt acts, and patients can also appease their voices by complying with less serious commands (a 'safety behaviour'). Our measure, developed from our

earlier work (Beck-Sander *et al*, 1997), recognises these subtleties and requires evidence not only from the client, but also from relatives or case managers. The rating of this scale was undertaken by three raters in the first instance to establish inter-rater reliability.

We are encouraged that the study also found (predicted) changes in power, distress/depression and omniscience (which were largely measured by self-report scales) and that these correlated significantly with our primary outcome, compliance; indeed, when power was controlled for in an analysis of covariance, the effect on compliance was rendered non-significant. This adds strength to our claim that compliance was genuinely changed and that the treatment effect was mediated by reduction in voice power; CTCH had broad effects on outcomes, but the absence of (self-reported) change in voice activity argues against the notion that patients' ratings were unreliable, and measures simply reflected the operation of a 'halo' effect in favour of CTCH across all measures.

Data obtained on prescription of neuroleptic and other drugs and provision of general mental health services during the course of the trial showed no difference between the groups and did not account for the effect of CTCH, but did underline their high level of service use linked to perceived risk. This suggests that the impact of CTCH we report here is unlikely to be accounted for by factors extraneous to the treatment. The pattern of neuroleptic use during the course of the trial showed no difference between the groups but did show a steady rise in neuroleptic prescription in the TAU group and a small reduction in the CTCH group. This suggests that concern about risk led to a raising of the dosage in those not receiving CTCH; this could reflect concern in clinicians as much as perceived benefit from CTCH.

The rise in neuroleptic use in TAU was correlated with reducing compliance; in CTCH the opposite was observed, *i.e.* reducing compliance was in line with reducing medication. There is a theoretical possibility that TAU participants were under-medicated and that the rise in medication prescription was responsible for the reduction in compliance (this could not account for the reduction in compliance in CTCH). This strikes us as unlikely for three reasons: first, both groups were receiving medication well in excess of *British National Formulary* (British Medical Association &

Royal Pharmaceutical Society of Great Britain, 2003) and other guidelines, including widespread use of atypicals and clozapine; second, at no point did dosage correlate with compliance, power or PANSS scores; and finally, using drug dosage as covariate did not affect the results. If the TAU group were under-medicated, this would serve to underestimate the effect size of CTCH, as compliance would be less likely to change over time. We believe that these differential changes in medication prescription reflect, as we indicate above, (understandable) clinician anxiety about this very high-risk group.

Nevertheless, it remains a possibility that non-specific aspects of the therapy were responsible for the effects. We believe, however, that the large correlation between changes in voice power and compliance by 6 months (0.63) strongly supports our contention that this aspect of the relationship with the voice (power) is the key independent variable. Whether it is cognitive therapy alone that brought about this change cannot be determined from these data, although we have clear evidence that the therapist adhered to protocol (see Method) and therefore we can be reasonably confident that the intervention itself was targeted at voice power and compliance.

The heterogeneity of the diagnosis of the sample is arguably a weakness. However, this was a pragmatic trial of the effect of CTCH on command hallucinations, and we decided to include all those who met the broad criteria for psychosis, irrespective of clinical diagnosis, to improve the generalisability of the findings. A second criticism is the apparently unequal numbers of those diagnosed with schizophrenia in each group (16 in TAU, 11 in CTCH; see Table 2). However, on the wider and widely used categorisation of schizophrenia and related disorders (ICD-10 F20-23, which includes schizoaffective disorders), this apparent difference becomes marginal in the other direction (16 in TAU, 15 in CTCH). Those with other diagnoses are also acceptably distributed, as are the scores on positive and negative symptoms. We would point out that the PANSS positive score was >20 in both groups (s.d.=3), indicating that our sample were indeed 'psychotic', notwithstanding the clinical diagnoses.

We feel the 'real world' relevance of the study is particularly strong. The sample as a whole was a severely distressed and a

generally high-risk group. Approximately 55% of those eligible took part (i.e. 38 out of 69), and 27% withdrew from the CTCH group, which is average for CBT in this population (Norman & Townsend, 1999; Durham *et al*, 2003). Given that this was a high-risk group, we looked at the reasons for patient withdrawal. We found, for example, that one person believed that the voice might harm or kill them for disclosing too much information; another feared that talking during therapy made the voices worse, and only continued on condition that it was the patient's decision how much to disclose about the voices.

The need for a further trial

The client group in this study – all experiencing command hallucinations and all having recently acted upon their commands – are typical of one of the highest-risk groups in psychiatry, who represent a major concern to their case managers, responsible medical officers and relatives, and particularly to themselves. This group is generally regarded as being resistant to treatment, whether with medication or cognitive-behavioural therapy (CBT) – 'conventional' CBT for psychosis is less effective with voices (Birchwood & Spencer, 2002) – and clinicians acknowledge equipoise in their management, as witnessed by the high level of referral to the trial. Our study showed that many clients felt themselves caught helplessly in a vortex of voice power, but found that CTCH gave them an opportunity to exert control by distancing themselves from their assumptions about voice power. For example, one client commented that, 'I know now that the voices can't hurt me – I feel that I am in control now. I still hear the voices but they are not as powerful.' Another client directly attributed his improvement to using 'all the techniques that she [S.B.] taught me; not only have the voices disappeared, but I am sleeping and eating properly now'. CTCH is therefore responding to a major gap in the treatment for people with ongoing, distressing voices, and deserves further evaluation.

This study was not definitive. It has suggested an effect size of major clinical significance, but because the sample size was small and the study was only conducted in one part of the country, there is a need to replicate it in a large-scale RCT incorporating different loci and different therapists, affording the opportunity to understand

for whom CTCH is most effective and how durable any effects might be. The durability question is of particular importance. There is a strand of psychiatric opinion that treatments for schizophrenia are only effective as long as they are active (McGlashan, 1988) and perhaps, therefore, a more theoretical and clinically relevant question might be 'how much further intervention is required to maintain the effect of treatment?'

Finally, Turkington *et al* (2003) observe in a recent editorial that current research into CBT for schizophrenia, although promising, is too imprecise, and that the way forward is to address specific questions, such as which are the active ingredients. They argue that trials with process measures 'will allow further clarification of the crucial elements of CBT for psychosis' (Turkington *et al*, 2003: p.98). Despite its limitations, we believe the present study is a step in this direction, in which the problem, the rationale, the intervention and the outcome are clearly specified and have a theoretical integrity and transparency, mediated through the process of the appraisal of voices' power.

ACKNOWLEDGEMENTS

The research undertaken for this study was supported by a grant from the Department of Health to P.T., M.B. and A.M. We are grateful to all the participants and the mental health staff who contributed to and supported the project in many ways. We would also like to acknowledge with thanks the advice of Professor Paul Chadwick.

REFERENCES

- Addington, D., Addington, J. & Maticka-Tyndale, E. (1993) Assessing depression in schizophrenia: the Calgary Depression Scale. *British Journal of Psychiatry*, **163** (suppl. 22), s39–s44.
- Beck-Sander, A., Birchwood, M. & Chadwick, P. (1997) Acting on command hallucinations: a cognitive approach. *British Journal of Clinical Psychology*, **36**, 139–148.
- Birchwood, M. & Spencer, E. (2002) Psychotherapies for schizophrenia. In *Schizophrenia: WPA Series in Evidence Based Psychiatry* (eds M. Maj & N. Sartorius), 2nd edn, pp. 147–241. Chichester: Wiley.
- Birchwood, M., Meaden, A., Trower, P., *et al* (2000) The power and omnipotence of voices: subordination and entrapment by voices and significant others. *Psychological Medicine*, **30**, 337–344.
- British Medical Association & Royal Pharmaceutical Society of Great Britain (2003) *British National Formulary* (September issue). London & Wallingford: BMJ Books & Pharmaceutical Press.
- Byrne, S., Trower, P., Birchwood, M., *et al* (2003) Command hallucinations: cognitive theory, therapy and

research. *Journal of Cognitive Psychotherapy: An International Quarterly*, **17**, 67–84.

Chadwick, P. & Birchwood, M. (1995) The omnipotence of voices II. The Beliefs About Voices Questionnaire (BAVQ). *British Journal of Psychiatry*, **166**, 773–776.

Chadwick, P., Lees, S. & Birchwood, M. (2000) The revised Beliefs About Voices Questionnaire (BAVQ-R). *British Journal of Psychiatry*, **177**, 229–232.

Cormac, I., Jones, C. & Campbell, C. (2002) Cognitive behaviour therapy for schizophrenia. *Cochrane Library*, issue 3. Oxford: Update Software.

Durham, R. C., Guthrie, M., Morton, R. V., et al (2003) Tayside–Fife clinical trial of cognitive–behavioural therapy for medication-resistant psychotic symptoms. *British Journal of Psychiatry*, **182**, 303–311.

Fleiss, J. (1981) *Statistical Methods for Rates and Proportions* (Wiley Series in Probability and Mathematical Statistics). Chichester: Wiley.

Gilbert, P. (1992) *Depression: the Evolution of Powerlessness*. Hove: Lawrence Erlbaum.

Gilbert, P., Birchwood, M., Gilbert, J., et al (2001) An exploration of evolved mental mechanisms for dominant and subordinate behaviour in relation to auditory hallucinations in schizophrenia and critical thoughts in depression. *Psychological Medicine*, **31**, 1117–1127.

Gould, R. A., Mueser, K. & Bolton, E. (2001) Cognitive behaviour therapy for schizophrenia: an effect size analysis. *Schizophrenia Research*, **48**, 335–342.

Haddock, G., Devane, S., Bradshaw, T., et al (2001) An investigation into the psychometric properties of the Cognitive Therapy Scale for Psychosis (CTS-Psy). *Behavioural and Cognitive Psychotherapy*, **29**, 221–234.

Haddock, G., McCarron, J., Tarrier, N., et al (1999) Scales to measure dimensions of hallucinations and delusions: the psychotic symptom rating scales (PSYRATS). *Psychological Medicine*, **39**, 879–889.

Jones, G., Huckle, P. & Tanaghaw, A. (1992) Command hallucinations, schizophrenia and sexual assaults. *Irish Journal of Psychological Medicine*, **9**, 47–49.

Kay, S. R., Fiszbein, A. & Opler, L. A. (1987) The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. *Schizophrenia Bulletin*, **13**, 261–269.

McGlashan, T. H. (1988) A selective review of North American long-term follow-up studies of schizophrenia. *Schizophrenia Bulletin*, **14**, 515–542.

CLINICAL IMPLICATIONS

- Cognitive therapy for command hallucinations (CTCH) has a comparatively large effect in reducing compliance with commands and delusional distress.
- CTCH significantly reduces the impact of ‘power’ beliefs which, according to our theory, have a causal role in compliance and are therefore a risk factor.
- CTCH is the first practical intervention that we know of that has a specific effect on compliance with command hallucinations.

LIMITATIONS

- This pilot study needs to be replicated in a full-scale randomised controlled trial.
- Although the treatment effect was mediated by a reduction in perceived voice power, we do not know if cognitive–behavioural therapy alone is responsible for this effect.
- Although the study has measured effects at 6 months’ follow-up, this is a relatively short time frame and we have no measure of longer-term effects.

PETER TROWER, PhD, MAX BIRCHWOOD, DSc, ALAN MEADEN, ClinPsyD, SARAH BYRNE, DCLinPsy, ANGELA NELSON, BA, KERRY ROSS, BSc, School of Psychology, University of Birmingham and Birmingham and Solihull Mental Health Trust, UK

Correspondence: Professor Max Birchwood, School of Psychology, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK. E-mail: M.J.Birchwood.20@Bham.ac.uk

(First received 12 August 2003, final revision 3 December 2003, accepted 15 December 2003)

Norman, R. & Townsend, L. (1999) Cognitive–behavioural therapy for psychosis: a status report. *Canadian Journal of Psychiatry*, **44**, 245–252.

Shawyer, F., Mackinnon, A., Farhall, J., et al (2003) Command hallucinations and violence: implications for detention and treatment. *Psychiatry, Psychology and Law*, **10**, 97–107.

Turkington, D., Kingdon, D. & Chadwick, P. (2003) Cognitive–behavioural therapy for schizophrenia: filling the therapeutic vacuum. *British Journal of Psychiatry*, **183**, 98–99.

World Health Organization (1992) *The ICD–10 Classification of Mental and Behavioural Disorders*. Geneva: WHO.