Dialogical therapy plus treatment as usual versus treatment as usual in adults with distressing and persistent auditory hallucinations in the UK (Talking with Voices-II): a randomised controlled trial

*Eleanor Longden, Alison Branitsky, Callum Glen, Verity Bell, Emmeline Joyce, Samantha E Bowe, Dirk Corstens, Rob Dudley, Amy Hardy, Bryony Sheaves, Richard Emsley, Melissa Pyle, Daniel Freeman, Wendy Jones, Anthony P Morrison, and the Talking with Voices Science Group**

Summary

Background Existing treatments for auditory hallucinations (hereafter voice hearing) among people with serious mental health problems (SMHPs; diagnoses include schizophrenia spectrum disorders, bipolar disorders, and personality disorders) show only modest clinical benefits, produce inconsistent long-term outcomes, and seldom prioritise personal recovery as a primary treatment goal. We aimed to investigate the effectiveness of Talking with Voices (TwV), a lived-experience informed intervention in which therapists engage in direct dialogue with voices, in improving both personal recovery and clinical outcomes for individuals experiencing persistent and distressing voice hearing.

Methods Talking with Voices-II was a rater-blinded, parallel-arm, multisite, randomised controlled trial to test the efficacy and underlying mechanisms of TwV therapy compared with treatment as usual (TAU) in community-based National Health Service mental health services across four UK sites. Participants with SMHPs who experienced distressing voices were randomly assigned (1:1), using random-sized blocks, stratified by site and presence of schizophrenia diagnosis, to receive either TwV plus TAU or TAU alone. Therapy was provided in up to 26 sessions over 8 months, and assessments were conducted at baseline, 8 months, and 14 months by assessors masked to treatment allocation. The primary outcome was the Questionnaire about the Process of Recovery score at 8 months, a self-reported measure of personal recovery in psychosis. Secondary outcomes were voice-related measures, self-beliefs, dissociation, and post-traumatic stress disorder (PTSD) symptoms. The primary analysis followed the intention-to-treat principle, and moderation and mediation were prespecified in the statistical analysis plan. The mediation analysis was carried out using the mediate command, utilising bootstrapping to obtain percentile CIs. The trial was prospectively registered with ISRCTN, ISRCTN15897915, and all research and clinical procedures were co-led by individuals with lived experience of psychosis.

Findings Between Sept 1, 2023, and June 30, 2025, we screened 1316 participants, of whom 309 consented, completed baseline assessments, and were randomly assigned to either TwV plus TAU (n=156) or TAU alone (n=153). The mean age of participants was 41·1 years (SD 13·1). 159 (52%) participants were men, 145 (47%) were women, and four (1%) were another gender, and one preferred not to say; 222 (72%) participants were White. At 8 months, personal recovery was significantly better in the TwV group than in the TAU group (adjusted difference 6·97 [95% CI 5·00–8·95]; p<0·0001; standardised effect size [SES] 0·63). Significant differences favouring TwV were found for 12 (80%) of 15 of secondary outcomes, unadjusted for multiple testing, including all voice-related measures, negative self-beliefs, and PTSD outcomes, with SES ranging from 0·18 to 0·86. There were no effects on dissociation-related measures or positive self-beliefs (SES ranging from <0·01 to 0·16). An effect was still evident on personal recovery at 14 months, although reduced (adjusted difference 3·55 [0·96–6·14]; p=0·0073; SES=0·32). 55 serious adverse events occurred in 36 participants in the TwV group, and 60 serious adverse events occurred in 45 participants in the TAU group, none of which were deemed definitely related to the therapy or research procedures.

Interpretation TwV significantly improves personal recovery, showing efficacy for individuals with SMHPs who experience distressing voices. Further research comparing TwV with an active control and evaluating its implementation in routine care settings is warranted.

Funding National Institute for Health and Care Research.

Copyright © 2026 The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY-NC-ND 4.0 license.

Lancet Psychiatry 2026

Published Online
November 10, 2026
https://doi.org/10.1016/S2215-0366(26)00330-5

See Online/Comment
https://doi.org/10.1016/S2215-0366(26)00345-7

*Group members listed at the end of the Article

Division of Psychology and Mental Health, School of Health Sciences, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre, The University of Manchester, Manchester, UK (E Longden PhD, A Branitsky PhD, Prof A P Morrison ClinPsyD); **Psychosis Research Unit, Greater Manchester Mental Health NHS Foundation Trust, Manchester, UK** (E Longden, A Branitsky, V Bell ClinPsyD, E Joyce ClinPsyD, S E Bowe ClinPsyD, M Pyle PhD, W Jones MSc, Prof A P Morrison); **Complex Trauma and Resilience Research Unit, Greater Manchester Mental Health NHS Foundation Trust, Manchester, UK** (E Longden, A Branitsky); **Department of Biostatistics and Health Informatics, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK** (C Glen PhD, R Emsley PhD); **GGZ Noord-Holland Noord, Texel/Den Helder, Netherlands** (D Corstens MD); **Gateshead Early Intervention in Psychosis Service, Northumbria, Tyne and Wear NHS Foundation Trust, Gateshead, UK** (Prof R Dudley PhD ClinPsyD); **Department of Psychology, University of York, York, UK** (Prof R Dudley); **Department of Psychology, Institute of Psychiatry, Psychology and**

Neuroscience, King's College London, London, UK (A Hardy PhD ClinPsyD)**; South London and Maudsley NHS Foundation Trust, London, UK** (A Hardy)**; Department of Experimental Psychology, University of Oxford, Oxford, UK** (B Sheaves ClinPsyD, Prof D Freeman PhD ClinPsyD)**; Oxford Health NHS Foundation Trust, Oxford, UK** (B Sheaves, Prof D Freeman)**; King's Clinical Trials Unit, King's College London, London, UK** (Prof R Emsley)

Correspondence to: Dr Eleanor Longden, Division of Psychology and Mental Health, School of Health Sciences, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre, The University of Manchester, Manchester, UK
eleanor.longden@manchester.ac.uk

Research in context

Evidence before this study

Over the past two decades, conceptualisations of voice hearing as an interpersonal and relational experience have attracted growing research attention. In response, a new wave of therapies has emerged, each centred upon understanding and transforming the relationship between hearer and voice. Talking with Voices (TwV) is one such approach, which specifically focusses on engaging with voices via direct dialogue. We searched PubMed from inception to May 6, 2026, for publications in any language using the search terms "talking with voices", "voice dialogue", "Maastricht interview", "experience focused counselling", "making sense of voices", and "TwV".

Previous research on dialogical approaches to voice hearing includes case examples, a concurrent multiple baseline design case series (n=15), a small randomised controlled trial (RCT; n=12), and the TwV feasibility RCT (n=50), all of which showed consistent signals of efficacy with no emergent safety concerns.

Several additional psychological interventions have been developed to reduce the severity of distressing voices. Meta-analysis evidence shows that cognitive behavioural therapy for psychosis, a treatment recommended by the National Institute for Health and Care Excellence, has moderate effects on voice hearing (d=0·46). Similarly, AVATAR therapy, in which participants engage in dialogue with a digital representation of their voice or voices, has been shown to achieve modest reductions in voice severity following treatment (Hedge's g=–0·40 [95% CI –0·54 to –0·25]). A meta-analysis of all psychological interventions that specifically target distressing voices found an effect size of 0·18 at end of treatment (95% CI –0·31 to –0·05) and 0·15 at follow-up (–0·27 to –0·03).

Added value of this study

This study represents the first fully powered multisite RCT of the TwV intervention, which is co-manualised and co-supervised by individuals with lived experience of psychosis. Compared with treatment as usual (TAU), findings indicate that TwV plus TAU leads to greater improvements in personal recovery among individuals with serious mental health problems (diagnoses include schizophrenia spectrum disorders, bipolar disorders, and personality disorders) who experience distressing voices at 8 months, with gains still evident, although reduced, at 14 months. These effects are mediated by changes in negative self-beliefs, beliefs about voices, and passive responses to voices, consistent with the proposed mechanisms. Additionally, compared with TAU alone, TwV was associated with greater reductions in voice severity and PTSD symptoms, as well as greater improvements in voice acceptance, at both 8-month and 14-month follow-ups. The observed reductions in voice severity represent, to our knowledge, the largest treatment effects reported for any RCT of a psychological intervention targeting voice hearing.

Implications of all the available evidence

The available evidence supports the use of TwV to improve personal recovery and reduce voice severity among individuals with serious mental health problems who experience distressing voices. The co-produced nature of the intervention might be central to its effectiveness, enhancing relevance, engagement, and acceptability. Further research is now warranted to examine TwV's effectiveness compared with an active control and implementation within routine care settings.

Introduction

Distressing auditory verbal hallucinations (hereafter voice hearing) are a debilitating feature of serious mental health problems (SMHPs), affecting 60–80% of individuals with non-affective psychosis as well as many with other SMHP diagnoses (SMHP diagnoses include schizophrenia spectrum disorders, bipolar disorders, and personality disorders).[1] The experience is associated with substantial functional impairment, suicidality, and increased risk of hospitalisation;[2] existing treatments show only modest effectiveness in reducing distress and supporting recovery. Specifically, estimates suggest traditional antipsychotic medications are ineffective for 25–30% of people who hear voices,[3] whereas cognitive behavioural therapy for psychosis (CBTp), the National Institute for Health and Care Excellence-recommended psychological intervention, yields clinically meaningful improvements in positive symptoms in approximately 25–53% of people with a diagnosis of schizophrenia,[4] with only small to moderate effect sizes for voices.[5] Moreover, although recovery is increasingly recognised as a central priority within mental health care, existing trials have largely prioritised symptom reduction as a primary outcome, with personal recovery rarely examined as a principal endpoint. Neither treatment strategy explicitly addresses the interpersonal dynamics between hearer and voice as an intervention target, despite growing consensus that these relationships might have a key role in maintaining distress and that directly modifying them can facilitate better patient outcomes.

In this regard, the Hearing Voices Movement, an influential network of people who hear voices and their allies, has long advocated that engaging therapeutically with voice content and meaning is central for promoting recovery.[6] Building on this perspective, a new wave of relationally informed interventions for voice hearing has begun to show promising outcomes among individuals with SMHP diagnoses.[7] Talking with Voices (TwV) extends this tradition by enabling therapists to engage directly in dialogue with an individual's voice or voices,

with the aim of fostering understanding and reconciliation within the hearer–voice relationship. In doing so, TwV targets various hypothesised psychological mechanisms, including dissociation, negative self-beliefs, relational dynamics with voices, and beliefs about and responses to voices. Specifically, dissociation was included in the trial's mechanistic model because it might link trauma to voices through alterations in consciousness, attention, and sense of self, and has been shown to mediate intrusive hallucinatory experiences more broadly; negative self-beliefs because these might often originate in trauma exposure while also being associated with more malevolent voice content, heightened distress, and poorer coping;[8–11] and voice relating on the basis that the ways voices are appraised and responded to (particularly fearful, submissive, or hostile patterns) might influence their maintenance and associated distress. Accordingly, we hypothesised that TwV would reduce dissociation and negative self-beliefs while promoting more adaptive relational dynamics, beliefs about and response to voices, and that changes in these processes would mediate the intervention's effects on personal recovery outcomes.

A pilot randomised controlled trial (RCT) involving 50 individuals with schizophrenia spectrum diagnoses who experienced distressing, personified voices that could be engaged in dialogue showed the feasibility and acceptability of the approach, with no emergent safety concerns.[12] Although not powered to detect clinical efficacy, the study showed signals of improvement in voice benevolence and personal recovery relative to treatment as usual (TAU). Collectively, these findings supported progression to a definitive, fully powered efficacy trial in a larger, transdiagnostic SMHP population.

We aimed to assess whether TwV added to TAU, compared with TAU alone, improved personal recovery in adults with SMHPs who experience persistent, distressing voices. Our primary hypothesis was that TwV plus TAU would lead to greater improvements in personal recovery at 8 months post-treatment compared with TAU alone. We further hypothesised that TwV plus TAU would be associated with improvements in broader indices of clinical recovery, and that these effects would be mediated by the specified mechanisms and maintained at the 14-month follow-up.

Method

Study design

The Talking with Voices-II trial was a rater-blinded, parallel-arm, multisite, randomised controlled trial assessing the efficacy and mechanisms of a psychological therapy targeting persistent, distressing voices (TwV) plus TAU, compared with TAU alone. The study was conducted in community, secondary adult mental health services across five National Health Service (NHS) [Trusts (Berkshire, Newcastle, Greater Manchester, Oxford, and South London) in the UK. Assessments and intervention delivery took place either on NHS premises, in voluntary sector organisations, or in participants' homes, with study procedures conducted either in person or remotely (via telephone or Microsoft Teams), depending on participant preference.

Independent monitoring and oversight were provided by a Trial Steering Committee, and ethics approval was received from Greater Manchester West NHS Research Ethics Committee (23/NW/0132). The study was prospectively registered on the ISRCTN registry (ISRCTN15897915), and reporting followed the CONSORT 2018 extension for social and psychological intervention trials.[13] The full protocol was published before the completion of data collection,[14] in accordance with SPIRIT guidelines,[15] and can be accessed online, along with the statistical analysis plan. All changes to the protocol were pre-approved by the sponsor, oversight committees and, where necessary, the Research Ethics Committee. No protocol changes occurred that altered the primary outcome, assessment timepoints, analysis plan, inclusion or exclusion criteria, or sample size; however, after trial commencement, one mechanistic outcome (Dissociative Experiences Scale-II) was updated to be included as a secondary clinical outcome. A substantial amendment to the protocol was made to include additional nested qualitative interviews with stakeholders, including trial participants.

For the **protocol and statistical analysis plan** see https://www.isrctn.com/ISRCTN15897915

The trial reflected the Hearing Voices Movement's foundational role in shaping the TwV intervention by prioritising patient and public involvement and engagement (PPIE) throughout design and delivery and embedding co-production at each relevant stage of implementation. PPIE included four members of the central research team who brought personal experience of psychosis (including the Chief Investigator [EL] and Trial Manager [AB]). Training for research staff and therapists was co-delivered by these two staff members alongside a clinical psychologist (SB) and psychiatrist (DC), and therapist supervision was co-delivered according to the same method, along with the development of the therapy manual. A researcher with lived experience of psychosis participated in Trial Management Group meetings as a standing contributor, and dedicated PPIE scrutiny of study procedures was facilitated by a lived-experience representative on the Trial Steering Committee. PPIE activity was tracked and actioned through standing agenda items at relevant meetings, and the Psychosis Research Unit's Service User Reference Group offered further consultation at relevant points in the study's design and implementation phases, including input into the development of participant-facing materials and the therapy manual.

Participants

The study population comprised users (aged ≥16 years) of NHS mental health services who heard persistent, distressing voices, who had heard voices for at least 1 year;

scored at least 1 on item 8 (amount of distress) of the Psychotic Symptom Rating Scales–Auditory Hallucinations Subscale (PSYRATS–AH);[16] were able to provide written informed consent; were actively seeking help in relation to distressing voices; were in contact with mental health services for at least 6 months; were willing and able to communicate with their voices, and to relay what the voices said to a therapist; and heard voices that were sufficiently personified to engage in dialogical work. Potential participants were excluded if they were at immediate risk of harm to self or others; were currently receiving structured, individual psychological therapy; did not speak English; had a primary diagnosis of alcohol or substance dependence or autism spectrum disorder; had a moderate or severe learning disability; had an organic cause for their voices; or were homeless or of no fixed abode. Details regarding establishing and maintaining interrater reliability for the PSYRATS–AH measure and operationalisation of the sufficient personification of the voices are given in the appendix (p 65).

See Online for appendix

Participants self-reported their gender and ethnicity via a standardised demographics questionnaire and were recruited through direct referrals from clinical teams, self-referrals, or approved research screening processes in line with local NHS Trust practice. Before the eligibility assessment, participants provided informed consent in either written or recorded verbal form.

Randomisation and masking

Following completion of their baseline assessment, participants were randomly assigned to either TwV plus TAU or TAU alone. Independent, concealed randomisation (1:1) was administered via a web-based system using random permuted blocks, stratified by site and diagnosis, by King's Clinical Trials Unit (UK Clinical Research Collaboration registration 053).

Masking was maintained for research assistants until outcome data had been collected for all participants. Participants, referring clinicians, therapists and their supervisors, and the central trial research team were unmasked; investigators remained masked to group-level summary data disaggregated by trial group throughout the study. Given the statistical analysis required to account for therapist-level clustering in the treatment group, the senior statistician (RE) was unmasked following database lock to oversee the analysis.

Across the trial, 49 masking breaks were reported (46 of these occurred before the 8-month follow-up; appendix p 66). Importantly, in every such case, the interviewer-administered outcome (PSYRATS–AH) was scored by an alternative, masked assessor. As a result, there were no unmasked assessment ratings.

Procedures

The trial followed a treatment manual developed and iteratively refined during the TwV-I pilot study.[17] TwV's primary goal is to improve the relationship between the person and their voice or voices and proceeds through four phases: (1) establishing a working alliance and exploring voice-hearing experiences, including discussions around normalisation and coping enhancement; (2) developing a personalised formulation that maps the core psychosocial conflicts associated with voice presence and content; (3) agreeing and pursuing goals for relational change including, where possible, direct dialogue facilitated by the therapist; and (4) consolidation and ending. Although direct dialogue might not be possible or necessary for every person who hears voices, the focus on the relationship with the voices and the psychological importance of what they say are common to all components.[17] In TwV, dialogue does not refer exclusively to the therapist or participant literally speaking to a voice, but rather how the intervention is dialogical in its broader focus on the voice–client relationship. Although direct dialogue provides a particularly distinctive way of exploring and potentially changing that relationship, it is TwV's broader relational focus (ie, exploring the voices' functional goals, meaning, intentions, and interpersonal dynamics) that is fundamental to the approach, not the completion of a prescribed number of discrete voice–client–therapist exchanges. Accordingly, TwV's other intervention stages and milestones should not be viewed as ancillary activities that are separate from the therapeutic mechanism, as all involve attending to and working with the relational dynamic and creating conditions in which a different relationship with the voices can develop, even when the voice is not directly addressed during the session (appendix p 114).

Therapy was delivered within an 8-month treatment window, allowing up to 26 individual 60-min therapy sessions, plus up to four optional booster sessions to consolidate therapeutic progress. Therapists held an accredited professional qualification, either as a clinical psychologist or as a CBT therapist with specialist psychosis experience. Supervision was provided in weekly groups by site and was jointly facilitated by a senior, experienced clinician (DC or SB) and a voice-hearer involved in the development and manualisation of the approach (EL or AB). Adherence to the therapy manual was monitored by therapists completing a bespoke TwV therapy session checklist after each session. The checklist comprised flexible milestones and therapeutic components derived from the manual. Adherence data were reviewed during trial delivery and milestones were discussed in supervision.

The control condition of TAU alone was provided by participants' existing mental health services. TAU encompassed the full range of multidisciplinary support typically available, including psychiatric medication, community-based health and social care, and access to rehabilitative and psychological services. Unless risk concerns arose requiring urgent attention, the TAU

condition did not entail any contact between the research team and participants' clinical teams. Similarly, referrers were not asked to restrict or modify any aspect of participants' treatment in either trial group during the study. Concomitant treatment endorsed by study participants across all timepoints is reported in the appendix (p 67).

Full details of screening and assessment procedures are given in the trial protocol.[14] Masked research assistants carried out assessments at baseline, 8 months post-randomisation (coinciding with the end of treatment), and a variable follow-up at 14 months post-randomisation. The last follow-up depended on when participants were recruited into the trial, resulting in a total follow-up period of either 8 months or 14 months; this approach was adopted to maximise the recruitment window and ensure the best value for money.

Outcomes

The primary outcome was personal recovery, assessed using the total score of the 15-item Questionnaire About the Process of Recovery (QPR)[18] at 8 months post-randomisation, with differences in mean scores serving as the analysis metric.

Secondary measures captured overall voice severity alongside broader dimensions of psychiatric distress and trauma-related sequelae. Total or subscale scores from each questionnaire or interview served as the measurement variables, with analyses examining differences in mean scores at both the primary and secondary endpoint.

Assessments relating to voice hearing and trauma were (1) voice severity (PSYRATS–AH), (2) Voices Acceptance and Action Scale (VAAS),[20] (3) PTSD Checklist for DSM-5 (PCL-5),[21] (4) Dissociative Experiences Scale (DES-II),[22] (5) Modified Trauma and Life Events Checklist (TALE),[23] (6) multi-modal hallucinations (PSYRATS-MMH), and (7) connections between adverse life events and voice hearing symptoms (TVAQ). Data for TALE, PSYRATS-MMH, and TVAQ are reported descriptively, as prespecified.

Assessments relating to the hypothesised mechanisms of action for TwV therapy were (1) the Brief Core Schema Scale (BCSS),[24] (2) the depersonalisation and derealisation subscale of the DES-II, (3) the Beliefs About Voices Questionnaire–Revised (BAVQ–R),[25] and (4) the Approve–Voices Questionnaire.[26]

Relevant demographic and clinical variables, including additional pharmacological and psychological treatments for mental health difficulties, were collected using a customised form, and a self-report, unpublished Learning From You evaluation questionnaire was also offered to participants on exit from the study (the latter

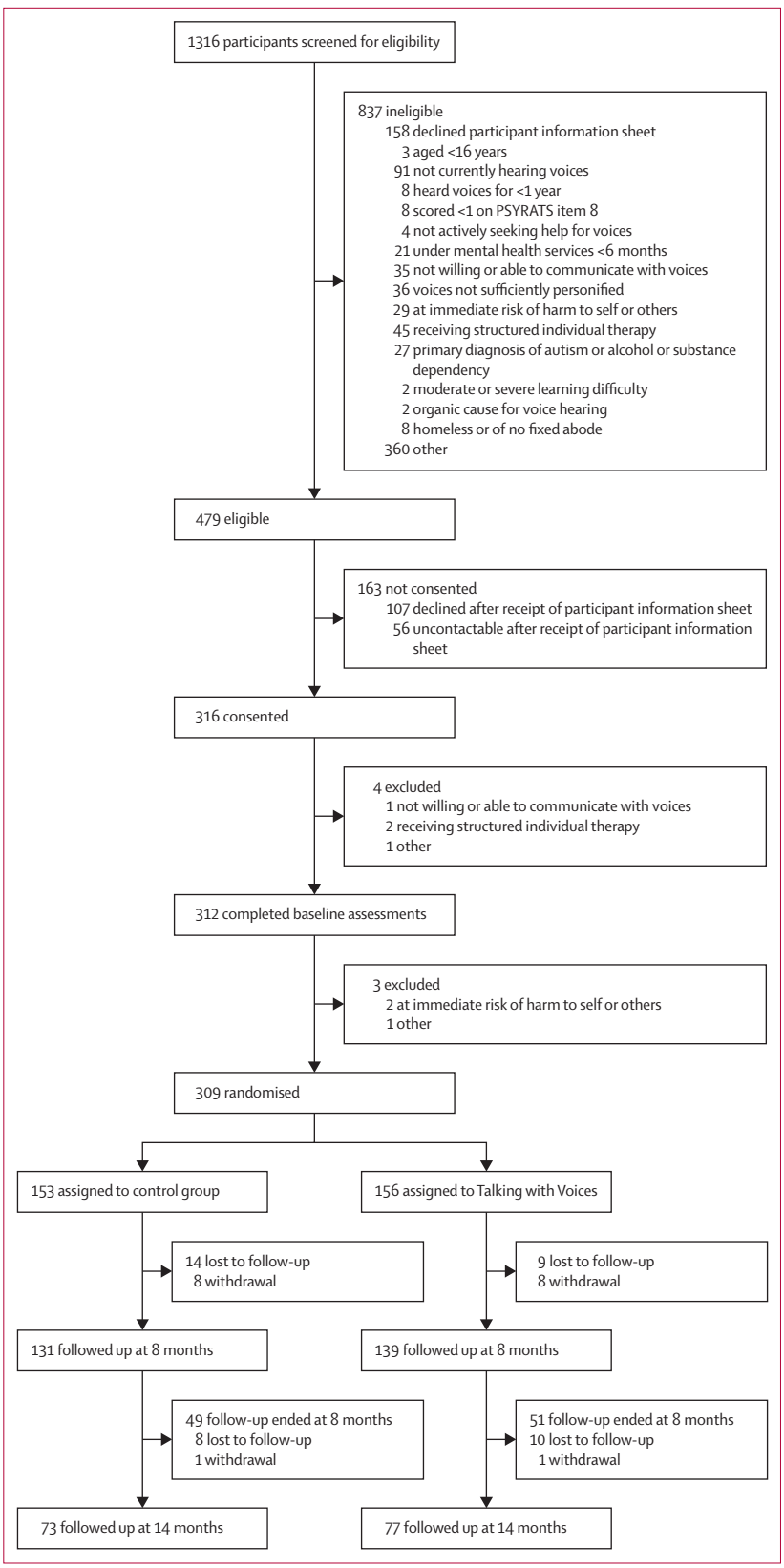


Figure 1: **Trial profile**
PSYRATS=Psychotic Symptom Rating Scale.

being intended to identify any unwanted or adverse effects of trial participation). Adverse events and serious adverse events were monitored throughout the study following Health Research Authority guidance for a trial that is not a clinical investigation of an investigational medicinal product, with data collected via spontaneous reporting from participants or their care teams, checking electronic patient records (EPRs) before assessment or intervention sessions, and systematic EPR screening at trial exit. Further details of the procedure for monitoring, classifying, and reporting serious adverse events are provided in the trial protocol.[14]

	Control (n=153)	Talking with Voices (n=156)	Total (n=309)
Age at date of consent, years			
Mean (SD)	40·50 (13·09)	41·73 (13·24)	41·12 (13·16)
Median (IQR)	40 (31–51)	40 (31–53)	40 (31–52)
Self-reported gender			
Male	84 (55%)	75 (48%)	159 (52%)
Female	66 (43%)	79 (51%)	145 (47%)
Non-binary	3 (2%)	0	3 (1%)
Other	0	1 (1%)	1 (<1%)
Prefer not to answer	0	1 (1%)	1 (<1%)
Highest level of education			
Never attended school	1 (1%)	0	1 (<1%)
Primary	11 (7%)	2 (1%)	13 (4%)
Secondary	45 (29%)	52 (33%)	97 (31%)
Further	56 (37%)	69 (44%)	125 (41%)
Higher	38 (25%)	32 (21%)	70 (23%)
Not reported	1 (1%)	1 (1%)	2 (1%)
Unknown	1 (1%)	0	1 (<1%)
Main employment status			
Full time	13 (9%)	22 (14%)	35 (11%)
Part time	7 (5%)	7 (5%)	14 (5%)
Voluntary	2 (1%)	6 (4%)	8 (3%)
Unemployed	124 (81%)	104 (67%)	228 (74%)
Retired	3 (2%)	9 (6%)	12 (4%)
Student	1 (1%)	5 (3%)	6 (2%)
Not available or not applicable	1 (1%)	2 (1%)	3 (1%)
Not reported	1 (1%)	1 (1%)	2 (1%)
Unknown	1 (1%)	0	1 (<1%)
Relationship status			
Single	109 (71%)	113 (72%)	222 (72%)
Married	13 (9%)	22 (14%)	35 (11%)
Divorced	13 (9%)	11 (7%)	24 (8%)
Separated	4 (3%)	3 (2%)	7 (2%)
Widowed	0	1 (1%)	1 (<1%)
Cohabiting	12 (8%)	5 (3%)	17 (6%)
Civil partnership	0	1 (1%)	1 (<1%)
Not reported	1 (1%)	0	1 (<1%)
Unknown	1 (1%)	0	1 (<1%)
Living arrangements			
With parents only	19 (12%)	15 (10%)	34 (11%)
With parents and siblings	11 (7%)	12 (8%)	23 (7%)
With partner	11 (7%)	21 (14%)	32 (10%)
Alone	61 (40%)	63 (40%)	124 (40%)
With friends	4 (3%)	0 (0%)	4 (1%)
With flatmates or housemates (not friends)	4 (3%)	7 (5%)	11 (4%)
With extended or other family	18 (12%)	20 (13%)	38 (12%)
Other	24 (16%)	18 (12%)	42 (14%)
Not reported	1 (1%)	0	1 (<1%)

(Table 1 continues on next page)

Choice of primary outcome

The QPR is a self-report instrument, co-developed with service users to capture personal recovery from psychosis, with items originally generated from qualitative interviews on this subject. The QPR has strong reliability, validity, and sensitivity to change, and has been adopted nationally as a patient-reported outcome measure for the assessment of early intervention for psychosis services. Personal recovery is consistently rated by patients as a higher priority than symptomatic change.[19] The minimum important difference for between-group comparisons is a 4-point change; thus, any between-group change greater than 4 points is considered clinically significant. The QPR is easily accessible, has been translated and validated in multiple languages including Spanish, French, Chinese, and Japanese, and is available for free use.

Statistical analysis

The trial used a partially nested design, with clustering arising from therapists in the intervention group; each control group participant was treated as a cluster of size 1. Assuming 14 therapists across the trial, an ICC of 0·02, an average of nine participants per therapist, and cluster size variation of nine (modelled on a Poisson process), 90% power to detect a between-group standardised effect size (SES) of 0·4 at 8 months required 252 participants with outcome data, based on a 5% two-sided significance level, a baseline-to-8-month score correlation of 0·4, and 1:1 allocation. Accounting for 15% attrition, 296 participants were required for recruitment. The target SES of 0·4 corresponds to a difference of 4·5 points on the primary outcome within the established minimum important difference of 4 points[27] using an SD of 11·5 derived from previous clinical trials of participants with SMHPs.

Descriptive statistics were used to summarise participant recruitment, attrition, and therapy completion rates. Primary analyses used an intention-to-treat (ITT) framework to target the treatment policy estimand. The ITT population comprised all randomly assigned participants, retained within their allocated group irrespective of any treatment switching. For the primary analysis, we fitted a linear mixed model to repeated QPR measurements obtained at the 8-month and 14-month timepoints, including the following fixed effects: treatment allocation, assessment timepoint, their interaction, recruitment site, presence of

schizophrenia spectrum diagnosis, and QPR score at baseline. Random intercepts were specified for participant, and for therapist in the TwV group only. The treatment policy estimand was operationalised as the covariate-adjusted mean difference between groups, derived separately for each timepoint with 8 months as the primary endpoint. Secondary analyses followed an analogous modelling strategy, using linear mixed models for continuous measures and logistic mixed models for binary endpoints. Missing outcome data were handled using full-information maximum likelihood estimation, assuming a missing-at-random mechanism conditional on all model covariates, with additional sensitivity analysis using a pattern-mixture model under missing not-at-random.

Causal mediation analysis using parametric regression was conducted to examine whether putative treatment-effect mechanisms at 8 months mediate any treatment effect on personal recovery and voice severity at 8 months, and separately at 14 months. We estimated the natural indirect effects of TwV through each hypothesised mediator separately controlling for site, diagnosis and baseline measures of outcome, and respective mediator. Mediators were only tested if they were significant when assessed as secondary outcomes at 8 months. No corrections were made for multiple comparisons in the mediation analysis.

Independent oversight was provided by the trial's Data-Monitoring and Ethics Committee (DMEC). Additional information on the analyses can be found in the statistical analysis plan, which was approved by the Trial Management Group, Trial Steering Committee, and DMEC, and is available via the trial's ISRCTN registry (ISRCTN15897915). Analyses were done in Stata (version 19).

Role of the funding source

The study protocol was reviewed and approved by the funder; however, the funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

We recruited participants from four NHS sites from Sept 1, 2023, to June 30, 2025, with the final assessment completed on March 13, 2026. We screened 1316 participants for eligibility, of whom 316 (24%) consented and were enrolled (figure 1). Seven participants were lost to follow-up after giving consent (because they did not wish to continue, their referral could not be progressed, or for unknown reasons), meaning that 309 participants were randomly allocated to the treatment (156 [50%]) and control (153 [50%]) groups. At baseline, 159 (52%) participants were male and 145 (47%) were female; three (1%) identified as non-binary, one as other, and one preferred not to identify. Mean age was 41 years (SD 13·2 [range 17–71]), and 222 (72%) were White or

	Control (n=153)	Talking with Voices (n=156)	Total (n=309)
(Continued from previous page)			
Ethnicity			
Asian or Asian British—Bangladeshi	0	3 (2%)	3 (1%)
Asian or Asian British—Indian	3 (2%)	5 (3%)	8 (3%)
Asian or Asian British—Pakistani	5 (3%)	2 (1%)	7 (2%)
Asian or Asian British—any other Asian background	0	4 (3%)	4 (1%)
Black or Black British—African	14 (9%)	12 (8%)	26 (8%)
Black or Black British—Caribbean	11 (7%)	8 (5%)	19 (6%)
Black or Black British—any other Black background	0	1 (1%)	1 (<1%)
Mixed—White and Asian	1 (1%)	1 (1%)	2 (1%)
Mixed—White and Black African	0	2 (1%)	2 (1%)
Mixed—White and Black Caribbean	2 (1%)	1 (1%)	3 (1%)
Mixed—any other mixed background	3 (2%)	1 (1%)	4 (1%)
White—British	108 (71%)	104 (67%)	212 (69%)
White—Irish	1 (1%)	4 (3%)	5 (2%)
White—any other White background	2 (1%)	3 (2%)	5 (2%)
Other ethnic group—any other ethnic group	2 (1%)	4 (3%)	6 (2%)
Not reported	1 (1%)	1 (1%)	2 (1%)
Religion or belief			
Atheism (no religious belief)	69 (45%)	62 (40%)	131 (42%)
Buddhism	1 (1%)	2 (1%)	3 (1%)
Christianity	46 (30%)	59 (38%)	105 (34%)
Islam	9 (6%)	10 (6%)	19 (6%)
Sikhism	2 (1%)	1 (1%)	3 (1%)
Hinduism	0	1 (1%)	1 (<1%)
Other	23 (15%)	18 (12%)	41 (13%)
Not reported	3 (2%)	2 (1%)	5 (2%)
Unknown	0	1 (1%)	1 (<1%)
Current psychiatric diagnosis			
Schizophrenia	50 (33%)	54 (35%)	104 (34%)
Psychosis	31 (20%)	27 (17%)	58 (19%)
Schizoaffective disorder	16 (11%)	15 (10%)	31 (10%)
PTSD	7 (5%)	8 (5%)	15 (5%)
Depression	5 (3%)	6 (4%)	11 (4%)
Bipolar disorder	8 (5%)	5 (3%)	13 (4%)
EUPD	9 (6%)	13 (8%)	22 (7%)
Other*	6 (4%)	7 (4%)	13 (4%)
Unknown	10 (6%)	8 (5%)	18 (6%)
Duration of voice hearing, years			
Mean (SD)	16·61 (13·59)	17·19 (13·36)	16·90 (13·46)
Median (IQR)	12·58 (6·00–25·00)	15·00 (6·00–25·50)	15·00 (6·00–25·00)
Number of voices heard in the last week			
Mean (SD)	6·70 (13·44)	4·80 (9·39)	5·73 (11·57)
Median (IQR)	3·00 (2·00–6·50)	3·00 (1·00–5·00)	3·00 (1·00–5·00)

Data are n (%), unless otherwise indicated. EUPD=emotionally unstable personality disorder. PTSD=post-traumatic stress disorder. *Other diagnoses in the control group were adjustment disorder, panic disorder, dissociative disorder, eating disorder, generalised anxiety disorder, and antisocial personality disorder, and in the Talking with Voices group were somatisation disorder, cyclothymia, dissociative disorder, generalised anxiety disorder, obsessive compulsive disorder, and personality disorder.

Table 1: **Baseline characteristics**

Caucasian, with Black or Black British (46 [15%]) as the second highest reported ethnicity (table 1).

At the end of treatment, 270 (87%) participants completed the 8-month assessment (139 [89%] in the TwV group and 131 [86%] in the control group). Most participants who were lost to follow-up between baseline and 8 months and between month 8 and month 14 either did not wish to complete assessments or could not be contacted (in the TvW group nine were lost to follow-up by month 8 and ten by month 14, and in the control group 14 were lost to follow-up by month 8 and eight by month 14; figure 1).

	Control	Talking with Voices	Adjusted mean difference (95% CI)	Standardised effect size	p value
Personal recovery (QPR score)*					
Baseline	28·72 (11·37)	27·47 (10·87)	..	..	..
8 months	30·47 (12·40)	36·60 (10·54)	6·97 (5·00 to 8·95)	0·63	<0·0001
14 months	33·81 (12·96)	34·72 (12·95)	3·55 (0·96 to 6·14)	0·32	0·0073
Voice severity (PSYRATS–AH score)					
Baseline	31·58 (4·02)	30·88 (4·27)	..	..	..
8 months	29·37 (6·95)	25·39 (7·65)	–3·56 (–5·40 to –1·72)	–0·86	0·0001
14 months	28·10 (8·37)	25·87 (7·74)	–3·30 (–6·03 to –0·58)	–0·79	0·017
Acceptance of voices (VAAS-12 score)*					
Baseline	36·61 (7·46)	36·60 (7·45)	..	..	..
8 months	38·12 (8·80)	43·09 (8·75)	4·58 (2·97 to 6·19)	0·62	<0·0001
14 months	39·33 (7·50)	41·17 (8·85)	3·34 (1·34 to 5·34)	0·45	0·0011
PTSD symptoms (PCL-5 score)					
Baseline	50·20 (16·59)	51·66 (14·73)	..	..	..
8 months	46·80 (18·94)	43·46 (17·43)	–4·64 (–8·65 to –0·62)	–0·30	0·024
14 months	43·58 (18·22)	42·31 (17·96)	–4·91 (–9·55 to –0·26)	–0·31	0·039
BCSS score					
Positive self					
Baseline	8·69 (6·11)	9·58 (6·39)	..	..	..
8 months	8·80 (5·88)	10·35 (6·26)	0·99 (–0·55 to 2·54)	0·16	0·21
14 months	9·48 (5·99)	10·52 (5·70)	1·48 (–0·28 to 3·25)	0·24	0·10
Negative self					
Baseline	11·33 (6·69)	11·45 (6·89)	..	..	..
8 months	10·47 (6·90)	7·74 (6·34)	–3·34 (–4·49 to –2·18)	–0·49	<0·0001
14 months	8·59 (6·35)	9·29 (6·63)	–1·64 (–3·25 to –0·02)	–0·24	0·047
DES-II score					
Total					
Baseline	41·91 (21·23)	43·33 (21·81)	..	..	..
8 months	39·43 (20·58)	39·93 (20·98)	–0·73 (–4·16 to 2·70)	–0·034	0·68
14 months	33·44 (21·02)	40·00 (19·68)	1·28 (–3·44 to 6·00)	0·060	0·60
Depersonalisation					
Baseline	43·57 (25·93)	43·62 (24·43)	..	..	..
8 months	40·03 (26·34)	39·17 (24·15)	–0·21 (–4·73 to 4·31)	–0·008	0·93
14 months	31·45 (24·05)	39·56 (22·53)	4·79 (–0·34 to 9·93)	0·19	0·067
Approve–Voices score					
Assertive					
Baseline	28·90 (13·93)	29·43 (13·18)	..	..	..
8 months	28·74 (12·32)	31·09 (12·68)	2·56 (0·27 to 4·86)	0·19	0·029
14 months	30·06 (14·33)	28·45 (13·98)	0·07 (–3·52 to 3·66)	0·005	0·97
Aggressive					
Baseline	27·07 (14·69)	26·64 (14·92)	..	..	..
8 months	26·05 (15·08)	20·11 (15·57)	–3·93 (–6·77 to –1·10)	–0·27	0·0066
14 months	24·42 (16·30)	20·08 (16·89)	–2·57 (–7·11 to 1·96)	–0·17	0·27
Passive					
Baseline	26·81 (13·19)	28·98 (12·88)	..	..	..
8 months	24·36 (13·85)	21·14 (14·08)	–4·43 (–7·74 to –1·12)	–0·34	0·0088
14 months	21·04 (13·66)	22·87 (13·97)	–1·52 (–5·40 to 2·36)	–0·12	0·44

(Table 2 continues on next page)

The median number of therapy sessions completed was 23 (mean 20·4 [SD 8·62; range 0–37]). Four participants assigned to the treatment group attended no sessions. There were 21 therapists, with a mean of 7 (SD 7; range 1–28) participants across therapists (appendix p 113).

In both the treatment and control groups, the personal recovery score increased (improved) from baseline to 8 months. The difference in QPR scores between the two groups at 8 months was statistically significant and higher in the TwV group (adjusted difference 6·97 [95% CI 5·00–8·95]; p≤0·0001; table 2). Using the compliance-adjusted analyses (complier average causal effect), the effect seen in the primary analysis was maintained with a slight increase (appendix p 88). The ICC for therapist clustering was <0·0001.

In both the control and treatment groups, we observed decreased voice severity and PTSD symptoms and increased voice acceptance scores, and we found significant between-group differences at 8 months in all secondary clinical outcomes favouring TwV, with the exception of dissociation, which showed no difference (table 2). Results were similar at 8 months (table 2) for nine of the 12 variables measuring postulated mechanisms, namely BCSS negative self subscale, and all BAVQ–R and Approve–Voices subscales (table 2).

The effect on the QPR at 14 months was also significantly improved in the TwV group with an SES of 0·32 (3·55 [95% CI 0·96–6·14; p=0·0073). A similar pattern of results was found at 14 months in voice severity and PTSD symptoms, voice acceptance, and mechanistic outcomes, namely the BCSS negative self subscale and the malevolence, benevolence, resistance, and engagement subscales of the BAVQ–R (table 2). Primary and secondary outcome analyses consistently favoured the therapy group (table 2). Forest plots of primary and secondary outcomes at 8 months and 14 months are shown in figure 2.

In the mediation analysis for QPR scores at 8 months, we found that the natural indirect effects were statistically significant for the BCSS negative self subscale, malevolence, benevolence, omnipotence, engagement, and resistance subscales of the BAVQ–R, and the passive subscale of the Approve–Voices questionnaire (table 3). A similar pattern of mediation was found for voice severity and for 14-month outcomes (appendix p 86). Planned sensitivity analyses, subgroup analyses, and complier average causal effect analyses are also reported in the appendix (pp 88–91).

Serious adverse events were balanced between groups, with 60 events in 45 participants in the TAU group, and 55 events in 36 participants in the TwV group (appendix pp 92–96). Most serious adverse events were classified as psychiatric disorders (42 [70%] in the TAU group and 33 [60%] in the TwV group). None of the serious adverse events were definitely related to either the study procedures or the study intervention. There were more adverse events reported in the TwV group (84 events in 50 participants *vs* 12 events in ten participants). Most adverse events were classified as psychiatric disorders (ten [83%] in the TAU group and 67 [80%] in the TwV group; appendix p 97). Descriptive data from TALE, PSYRATS-MMH, TVAQ, and our Learning From You measures are reported as was prespecified (appendix pp 11–65, 101–108).

Discussion

The trial confirmed our primary hypothesis of clinical efficacy, showing a significantly greater improvement in the intervention group than in the control group by the end of therapy for personal recovery (SES 0·63); this difference is clinically significant, exceeding the minimum important difference for between-group comparisons with the QPR (adjusted difference 6·97 [95% CI 5·00–8·95]). Results for voice severity at month 8 similarly showed substantial effects (SES 0·86). Although not directly comparable, the treatment effect on voice severity in this trial is larger than the mean effect reported for voice-related outcomes in meta-analyses of both CBTp (d=0·46)[28] and AVATAR therapy (a therapeutic approach in which participants engage in dialogue with a digital representation of their voice or voices; Hedges' g=–0·40 [95% CI –0·54 to –0·25).[29] To our knowledge, this is the first powered RCT to investigate a dialogical intervention such as TwV, examining both its efficacy and underlying mechanisms in participants with co-occurring SMHP diagnoses and distressing voice hearing. Lived-experience leadership and collaboration was integral to all aspects of the trial. Importantly, TwV therapy was also feasible to deliver, with low attrition, acceptable to participants, and not associated with any therapy-related adverse events.

The 14-month findings across four of five clinical outcomes remained significant, albeit reduced, indicating that therapeutic gains achieved through TwV were still evident at follow-up. This finding is particularly noteworthy given the planned reduced sample size at 14 months (n=150, compared with n=276 at the end of treatment). The meta-analysis[31] of targeted psychological intervention for auditory hallucinations found an SMD of 0·15 for voice severity at follow-up—our effect size for this outcome was notably larger at follow-up (0·79; table 2).

The mediation analyses were consistent with the hypothesised mechanisms, suggesting that the therapy improves both personal recovery and voice severity by

	Control	Talking with Voices	Adjusted mean difference (95% CI)	Standardised effect size	p value
(Continued from previous page)					
BAVQ–R score*					
Malevolence					
Baseline	11·90 (4·54)	11·66 (4·37)	..	..	..
8 months	11·08 (5·10)	7·44 (5·68)	–3·37 (–4·61 to –2·12)	–0·76	<0·0001
14 months	10·29 (5·23)	8·40 (5·82)	–2·35 (–3·83 to –0·87)	–0·53	0·0018
Benevolence					
Baseline	3·63 (4·54)	3·92 (4·48)	..	..	..
8 months	3·25 (4·36)	6·85 (5·21)	3·17 (1·99 to 4·35)	0·70	<0·0001
14 months	3·31 (5·03)	5·59 (4·78)	2·34 (1·14 to3·53)	0·52	0·0001
Omnipotence					
Baseline	12·13 (3·91)	11·88 (3·54)	..	..	..
8 months	11·20 (4·25)	9·50 (4·33)	–1·61 (–2·42 to –0·80)	–0·43	0·0001
14 months	10·07 (4·91)	10·22 (4·63)	–0·70 (–2·06 to 0·65)	–0·19	0·31
Resistance					
Baseline	19·69 (4·79)	19·43 (5·12)	..	..	..
8 months	18·45 (5·58)	14·83 (7·42)	–3·45 (–5·06 to –1·84)	–0·70	<0·0001
14 months	18·04 (6·07)	16·58 (7·23)	–2·32 (–4·16 to –0·48)	–0·47	0·013
Engagement					
Baseline	5·31 (5·59)	5·44 (4·96)	..	..	..
8 months	4·97 (5·93)	8·16 (6·55)	3·03 (1·64 to 4·42)	0·58	<0·0001
14 months	5·00 (6·43)	6·71 (6·48)	2·02 (0·51 to 3·52)	0·38	0·0087

Data are mean (SD), unless otherwise indicated. Estimates are adjusted for recruitment site, schizophrenia spectrum diagnosis, and outcome score at baseline. BAVQ–R=Beliefs About Voices Questionnaire–Revised. BCSS=Brief Core Schema Scale. DES-II=Dissociative Experiences Scale. PCL-5=PTSD Checklist for DSM-5. PSYRATS–AH=Psychotic Symptom Rating Scale–Auditory Hallucinations Subscale. QPR=15-item Questionnaire About the Process of Recovery. VAAS-12=Voices Acceptance and Action Scale. *Higher scores indicate better outcomes.

Table 2: **Primary and secondary outcomes**

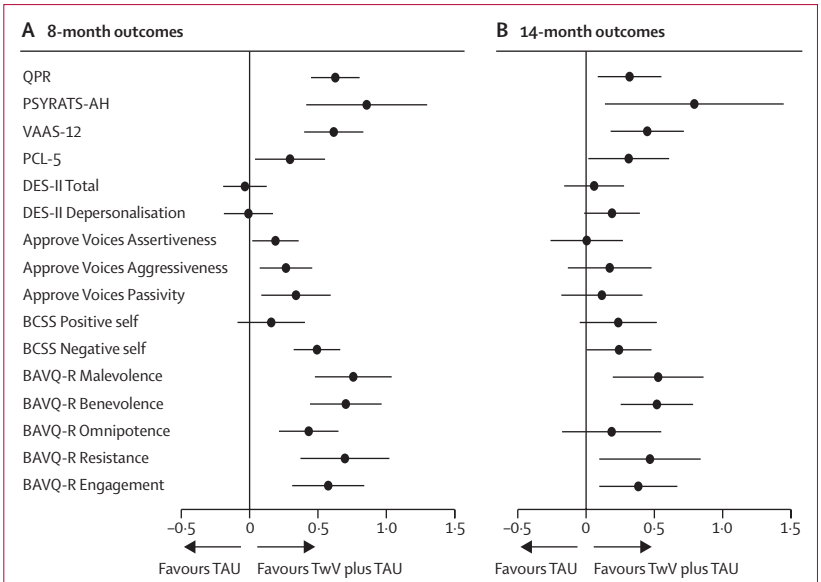


Figure 2: **Forest plots of primary and secondary outcomes at 8 months and 14 months**
BAVQ–R=Beliefs About Voices Questionnaire–Revised. BCSS=Brief Core Schema Scale. DES-II=Dissociative Experiences Scale. PCL-5=PTSD Checklist for DSM-5. PSYRATS–AH=Psychotic Symptom Rating Scale–Auditory Hallucinations Subscale. QPR=15-item Questionnaire About the Process of Recovery. TAU=treatment as usual. TwV=Talking with Voices. VAAS-12=Voices Acceptance and Action Scale.

	Natural indirect effect		Natural direct effect		Total effect		Percentage of total effect
	Effect (95% CI)	p value	Effect (95% CI)	p value	Effect (95% CI)	p value	
BCSS							
Negative self	2·98 (1·48 to 4·82)	0·0007	3·56 (1·29 to 6·02)	0·0038	6·54 (4·19 to 8·94)	<0·0001	45·56
BAVQ–R							
Malevolence	1·96 (0·87 to 3·47)	0·0028	4·11 (1·77 to 6·52)	0·0006	6·07 (3·81 to 8·41)	<0·0001	32·29
Benevolence	2·17 (0·94 to 3·57)	0·0014	3·93 (1·49 to 6·44)	0·0018	6·10 (3·82 to 8·42)	<0·0001	35·57
Omnipotence	1·92 (0·80 to 3·18)	0·0016	4·13 (2·12 to 5·92)	<0·0001	6·05 (3·77 to 8·36)	<0·0001	31·74
Resistance	1·30 (0·48 to 2·39)	0·0080	4·87 (2·58 to 7·15)	<0·0001	6·17 (3·81 to 8·34)	<0·0001	21·07
Engagement	1·20 (0·21 to 2·52)	0·042	5·10 (2·88 to 7·59)	<0·0001	6·30 (4·17 to 8·61)	<0·0001	19·05
Approve–Voices							
Assertiveness	0·19 (–0·10 to 0·89)	0·45	5·92 (3·58 to 8·17)	<0·0001	6·11 (3·75 to 8·35)	<0·0001	3·11
Aggression	0·86 (0·13 to 1·94)	0·061	5·26 (2·90 to 7·46)	<0·0001	6·10 (3·69 to 8·35)	<0·0001	14·10
Passive	1·34 (0·38 to 2·64)	0·019	4·99 (2·92 to 7·01)	<0·0001	6·34 (3·95 to 8·57)	<0·0001	21·14

Each analysis was conducted independently of the other mediators. BAVQ–R=Beliefs About Voices Questionnaire–Revised. BCSS=Brief Core Schema Scale. QPR= 15-item Questionnaire About the Process of Recovery. Mediators that were significant as secondary outcomes at 8 months were tested. No corrections were made for multiple comparisons in the mediation analysis.

Table 3: **Mediation analysis results for personal recovery (QPR) at 8 months**

changing negative self-beliefs, voice-related appraisals, and patterns of relating. It is also important to contextualise these findings within the characteristics of the trial population. Specifically, participants presented with persistent SMHPs (193 [62%] of 309 had schizophrenia spectrum diagnoses) and reported unremitting and distressing voice hearing lasting a minimum of 12 months before enrolment (mean duration 16·90 years [SD 13·46]). Notably, baseline voice severity scores were similar to those reported in trials of AVATAR therapy,[29] supporting the representativeness of the sample and the generalisability of the findings.

Several limitations should be considered when interpreting these findings. The absence of an active comparator, such as CBTp or supportive counselling, in combination with the fact that participants were not masked to allocation, might create an expectation bias. The absence of an active comparator also raises the possibility that differences were attributable to non-specific effects from 8 months of intensive intervention.

The lack of a validated competency scale precluded formal competency assessment, and recordings of therapy sessions were instead used during supervision to provide informal feedback on therapist competency. Therefore, control of therapist adherence represents a potential source of bias and variation in therapist skills and adherence might have contributed to the observed effects. Decisions regarding eligibility in relation to the sufficient personification of voices were based on a structured review of information; as such, no inter-rater reliability data were available for this determination (appendix p 65).

107 (22%) of 479 eligible individuals declined to participate following receipt of the participant information sheet, which might indicate that the therapeutic aim of improving one's relationship with voices was unappealing to a subset of people who hear voices. By contrast, comparatively few individuals were excluded because they were unable or unwilling for a therapist to engage in dialogue with their voices (n=35 [4%]), or because their voices were insufficiently personified to permit dialogue (n=36 [4%]), suggesting that recruitment difficulties might have reflected challenges regarding readiness for the intervention approach, rather than limitations in the applicability of TwV to the wider population of people who hear voices. This represents a potential challenge for future implementation.

Although the sample included 46 (15%) Black or Black British participants and 22 (7%) Asian or Asian British participants, the predominantly White or Caucasian participant composition limits the generalisability of the findings to more ethnically diverse populations. The high rate of ineligible screened participants largely reflects the fact that, at several sites, entire community mental health team caseloads were screened rather than pre-selected individuals. Our sample is likely to be representative of the majority of people who hear voices within psychiatric services in the UK. In terms of dialogical voices specifically, large phenomenological surveys indicate that most people who hear voices report some form of relationship with their voices[30] and that these voices possess distinct characterful qualities. Similarly, distress is almost universal among people who hear voices receiving psychiatric care, and a requirement for target symptoms to be causing distress is standard across trials of psychological therapies. Therefore, our results are likely to be generalisable to most people who hear voices receiving mental health services.

The therapy was delivered by skilled practitioners within a framework of intensive joint supervision

(provided by both experienced clinicians and lived-experience supervisors), which might limit the generalisability of findings to NHS mental health services as currently configured. Similarly, the prevailing culture within routine practice, which tends to prioritise the eradication of voices over fostering a more peaceful coexistence with them, might present a structural barrier to the wider implementation of TwV's Hearing Voices Movement-informed approach, and the feasibility of implementation within such a context warrants careful consideration and further research. The reduced numbers of participants at follow-up (due to funding constraints that meant those recruited in the last 6 months of the trial were not followed up at 14 months) resulted in reduced statistical power at this timepoint. However, this was a pre-planned feature of our trial design, which applied systematically to both groups, so should not represent a source of bias.

With respect to attrition, our dropout rate was low, particularly given the study population (the Cochrane Risk of Bias tool[32] explicitly notes that in trials involving schizophrenia-spectrum populations, ongoing symptoms commonly increase the likelihood of dropout). At our primary endpoint, there were only 16 (5%) formal withdrawals and 23 (7%) participants lost to follow-up, giving an overall attrition rate of 13%, which is lower than the 15% anticipated in our sample size calculation.

The findings of this trial carry several important clinical implications. Most directly, our findings expand the range of evidence-based options available to people experiencing distressing voices, offering therapeutic pathways oriented towards outcomes that have received comparatively little attention in the literature, such as improved relationships between people who hear voices and their voices, as well as facilitating personal recovery and reducing voice severity. The trial also shows the value of embedding lived experience within clinical supervision and approaching voices as a transdiagnostic experience, and points to the potential for survivor-informed models to meaningfully shape therapeutic delivery.

The findings also indicate several important directions for future research. Given the strong effects observed, and in view of the challenges in aligning TwV's approach with a prevailing biomedical culture of mental health services, as well as concerns about the appropriateness of engaging with voices, implementation research is a clear priority to investigate the potential for translation to routine care. This research might include an investigation of factors that influence people who hear voices' readiness to engage with the therapy, possibly incorporating testimonials from those who have received and benefited from TwV, to facilitate uptake among those reluctant to consider the therapy. Further research could also examine the cultural adaptations required to improve accessibility and inclusion and the specific training needed to support clinical teams in delivering the approach.

More broadly, research examining what works and for whom among people who hear distressing voices is needed, with the aim of identifying which evidence-based approaches are best suited to a given individual considering their circumstances and desired outcomes. Head-to-head comparisons with active controls are also important.

In summary, in this efficacy trial, TwV plus TAU was superior to TAU in facilitating personal recovery and reducing voice severity in a transdiagnostic sample experiencing persistent and distressing voices. These outcomes were mediated by changes in beliefs about voices, negative beliefs about self, and relational dynamics between hearer and voice or voices.

The Talking with Voices Science Group

Natalie Barnes, Poppy Brown, Lauren Cash, Grace Gowers, Amy Langman, Rachel Manser, Charis Palmer, and Jerica Radez (Oxford Health NHS Foundation Trust); Vicky Brooks, Beth Cooper, Tara Doran, Lucy Loftus, Emma De Paola, Cyd Ross, Jake Smart, Hannah Sowery, and Kim Towey-Swift (Greater Manchester Mental Health NHS Foundation Trust); Tenniel Garwes, Sam Gleaden, Geetanjali Kumar, Catherine Lankester, Leila Lawton, Kathryn Taylor, and Andrew Watson (South London and Maudsley NHS Foundation Trust); Jane Mitchell, Pamela O'Hara, Lucy Patterson, Louise Prentice, and Alexandra Williams (Cumbria, Northumberland, Tyne and Wear NHS Foundation Trust).

Contributors

EL, AB, CG, SEB, DC, RD, AH, BS, RE, MP, DF, WJ, and APM contributed to study conceptualisation. EL, AB, VB, EJ, SEB, DC, RD, AB, BS, and APM contributed to data curation. CG and RE contributed to formal analysis. EL, AB, CG, SEB, DC, RD, AH, BS, RE, MP, DF, WJ, and APM contributed to funding acquisition. EL, AB, VB, EJ, SEB, DC, RD, AB, BS, and APM contributed to investigation. EL, AB, CG, SEB, DC, RD, AH, BS, RE, MP, DF, and APM contributed to methodology. AB, VB, and EJ contributed to project administration. EL, AB, VB, EJ, SEB, DC, RD, AH, BS, and APM supervised the study. CG and RE contributed to data validation and visualisation. EL, AB, CG, and APM wrote the original draft. VB, EJ, SEB, DC, RD, AH, BS, RE, MP, DF, WJ, and APM contributed to manuscript review and editing. RE and CG verified the data and had access to raw data. EL had final responsibility for the decision to submit for publication.

Declaration of interests

EL has received payment for delivering presentations and workshops on both the TwV and Hearing Voices Movement approach. AB has received payment for delivering presentations and workshops on both the TwV and Hearing Voices Movement approach and is a trustee of the English Hearing Voices Network. DC has received payment for delivering presentations and workshops on both the TwV and Hearing Voices Movement approach and book publications on the Hearing Voices Movement. RD receives payment for workshops on working with unusual experiences like hallucinations, receives grant income for studies of psychological therapies for psychosis, and works in the NHS delivering psychological therapies. AH has received payment for training and supervision in trauma therapies for psychosis. All other authors declare no competing interests.

Data sharing

De-identified participant data will be available in anonymised form from the corresponding author on reasonable request, subject to review and contract with Greater Manchester Mental Health NHS Foundation Trust.

Acknowledgments

The authors gratefully acknowledge all participants who agreed to take part in the trial. We are grateful to the Psychosis Research Unit Service User Reference group for their collaboration on the design of the study and contribution to the development of study materials. We are grateful to our independent Trial Steering Committee (Andrew Gumley,

Alison Brabban, Obi Ukoumunne, and Stephanie Allan) and our independent Data Monitoring Committee (Matthias Schwannauer, Anne Cooke, and Anna-Marie Bibby-Jones) for providing oversight on the trial. We are grateful for the many trial therapists, research workers, and network staff who supported the study. The study was facilitated by the Greater Manchester Local Clinical Research Network. This study was supported by the UK Clinical Research Collaboration-registered King's Clinical Trials Unit at King's Health Partners. The study was funded by the National Institute of Health Research Efficacy and Mechanism Evaluation Programme (NIHR151223). RE was supported by NIHR Research Professorship (NIHR300051) and the National Institute for Health and Care Research (NIHR) Biomedical Research Centre–Maudsley. The views and opinions expressed are those of the authors and do not necessarily reflect those of the Efficacy and Mechanism Evaluation Programme, the National Institute for Health Research, NHS, or the Department of Health and Social Care.

References

1 Lim A, Hoek HW, Deen ML, et al. Prevalence and classification of hallucinations in multiple sensory modalities in schizophrenia spectrum disorders. *Schizophr Res* 2016; **176:** 493–99.

2 Waters F, Fernyhough C. Hallucinations: a systematic review of points of similarity and difference across diagnostic classes. *Schizophr Bull* 2017; **43:** 32–43.

3 Shergill SS, Murray RM, McGuire PK. Auditory hallucinations: a review of psychological treatments. *Schizophr Res* 1998; **32:** 137–50.

4 Bighelli I, Huhn M, Schneider-Thoma J, et al. Response rates in patients with schizophrenia and positive symptoms receiving cognitive behavioural therapy: a systematic review and single-group meta-analysis. *BMC Psychiatry* 2018; **18:** 380.

5 Turner DT, Burger S, Smit F, Valmaggia LR, van der Gaag M. What constitutes sufficient evidence for case formulation-driven CBT for psychosis? Cumulative meta-analysis of the effect on hallucinations and delusions. *Schizophr Bull* 2020; **46:** 1072–85.

6 Longden E, Corstens D. Making sense of voices: perspectives from the Hearing Voices Movement. In: Berry K, Bucci S, Danquah AN, eds. Attachment theory and psychosis: current perspectives and future directions. London: Routledge, 2020: 223–36.

7 Dellazizzo L, Giguère S, Léveillé N, Potvin S, Dumais A. A systematic review of relational-based therapies for the treatment of auditory hallucinations in patients with psychotic disorders. *Psychol Med* 2022; **52:** 2001–08.

8 Andrew EM, Gray NS, Snowden RJ. The relationship between trauma and beliefs about hearing voices: a study of psychiatric and non-psychiatric voice hearers. *Psychol Med* 2008; **38:** 1409–17.

9 Bloomfield MAP, Chang T, Woodl MJ, et al. Psychological processes mediating the association between developmental trauma and specific psychotic symptoms in adults: a systematic review and meta-analysis. *World Psychiatry* 2021; **20:** 107–23.

10. Hardy A. Pathways from trauma to psychotic experiences: a theoretically informed model of posttraumatic stress in psychosis. *Front Psychol* 2017; **8**.

11 Alameda L, Rodriguez V, Carr E, et al. A systematic review on mediators between adversity and psychosis: potential targets for treatment. *Psychol Med* 2020; **50:** 1966–76.

12 Longden E, Corstens D, Bowe S, et al. A psychological intervention for engaging dialogically with auditory hallucinations (Talking with Voices): a single-site, randomised controlled feasibility trial. *Schizophr Res* 2022; **250:** 172–79.

13 Grant S, Mayo-Wilson E, Montgomery P, et al. CONSORT-SPI 2018 explanation and elaboration: guidance for reporting social and psychological intervention trials. *Trials* 2018; **19:** 406.

14 Longden E, Bell V, Joyce E, et al. Multisite randomised controlled trial of a novel dialogical therapy in comparison to treatment as usual in adults with distressing and persistent auditory hallucinations: study protocol for the Talking with Voices (TWV-II) trial. *Trials* 2025; **26:** 405.

15 Chan AW, Tetzlaff JM, Gøtzsche PC, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. *BMJ* 2013; **346:** e7586.

16 Haddock G, McCarron J, Tarrier N, Faragher EB. Scales to measure dimensions of hallucinations and delusions: the psychotic symptom rating scales (PSYRATS). *Psychol Med* 1999; **29:** 879–89.

17 Longden E, Corstens D, Morrison AP, et al. A treatment protocol to guide the delivery of dialogical engagement with auditory hallucinations: experience from the Talking with Voices pilot trial. *Psychol Psychother* 2021; **94:** 558–72.

18 Law H, Neil ST, Dunn G, Morrison AP. Psychometric properties of the questionnaire about the process of recovery (QPR). *Schizophr Res* 2014; **156:** 184–89.

19 Law H, Morrison AP. Recovery in psychosis: a Delphi study with experts by experience. *Schizophr Bull* 2014; **40:** 1347–55.

20 Brockman R, Kiernan M, Murrell E. Psychometric properties of two brief versions of the Voices Acceptance and Action Scale (VAAS): implications for the second-wave and third-wave behavioural and cognitive approaches to auditory hallucinations. *Clin Psychol Psychother* 2015; **22:** 450–59.

21 Blevins CA, Weathers FW, Davis MT, Witte TK, Domino JL. The posttraumatic stress disorder checklist for DSM-5 (PCL-5): development and initial psychometric evaluation. *J Trauma Stress* 2015; **28:** 489–98.

22 Carlson E, Putnam F. An update on the Dissociative Experiences Scale. *Dissociation* 1993; **6:** 16–27.

23 Carr S, Hardy A, Fornells-Ambrojo M. The Trauma and Life Events (TALE) checklist: development of a tool for improving routine screening in people with psychosis. *Eur J Psychotraumatol* 2018; **9:** 1512265.

24 Fowler D, Freeman D, Smith B, et al. The Brief Core Schema Scales (BCSS): psychometric properties and associations with paranoia and grandiosity in non-clinical and psychosis samples. *Psychol Med* 2006; **36:** 749–59.

25 Chadwick P, Lees S, Birchwood M. The revised Beliefs About Voices Questionnaire (BAVQ–R). *Br J Psychiatry* 2000; **177:** 229–32.

26 Hayward M, Schlier B, Strauss C, Rammou A, Lincoln T. Construction and validation of the Approve questionnaires—measures of relating to voices and other people. *Schizophr Res* 2020; **220:** 254–60.

27 Dehmahdi N, Law H, Pyle M, et al. Estimating the minimum important difference for the questionnaire about the Process of Recovery (QPR): an anchor-based approach. *Psychosis* 2021; **13:** 220–30.

28 van der Gaag M, Valmaggia LR, Smit F. The effects of individually tailored formulation-based cognitive behavioural therapy in auditory hallucinations and delusions: a meta-analysis. *Schizophr Res* 2014; **156:** 30–37.

29 Opper F, Henges S, Weinstein P, et al. Audio Visual Assisted Therapy Aid For Refractory Auditory Hallucinations (AVATAR) therapy for voice hearers: systematic review and meta-analysis. *BJPsych Open* 2026; **12:** e104.

30 McCarthy-Jones S, Trauer T, Mackinnon A, Sims E, Thomas N, Copolov DL. A new phenomenological survey of auditory hallucinations: evidence for subtypes and implications for theory and practice. *Schizophr Bull* 2014; **40:** 231–35.

31 Fässler L, Koop S, Opper F, et al. Psychological interventions targeting auditory hallucinations in persons with psychotic disorders: a systematic review and meta-analysis. *JAMA Psychiatry* 2026; **83:** 1007–19.

32 Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019; **366:** l4898.