



## UvA-DARE (Digital Academic Repository)

### Effectiveness of Imagery Rescripting for Trauma-Affected Voice Hearers

*An Open Trial*

Strachan, L.; McEvoy, P.; Arntz, A.; Steel, C.; Paulik, G.

**DOI**

[10.1002/cpp.3052](https://doi.org/10.1002/cpp.3052)

**Publication date**

2024

**Document Version**

Final published version

**Published in**

Clinical Psychology and Psychotherapy

**License**

CC BY

[Link to publication](#)

**Citation for published version (APA):**

Strachan, L., McEvoy, P., Arntz, A., Steel, C., & Paulik, G. (2024). Effectiveness of Imagery Rescripting for Trauma-Affected Voice Hearers: An Open Trial. *Clinical Psychology and Psychotherapy*, 31(5), Article e3052. <https://doi.org/10.1002/cpp.3052>

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

## RESEARCH ARTICLE OPEN ACCESS

# Effectiveness of Imagery Rescripting for Trauma-Affected Voice Hearers: An Open Trial

Laura Strachan<sup>1</sup> | Peter McEvoy<sup>1</sup>  | Arnoud Arntz<sup>2</sup> | Craig Steel<sup>3,4</sup> | Georgie Paulik<sup>1,5,6</sup> 

<sup>1</sup>School of Psychology, Curtin University, Curtin, Western Australia, Australia | <sup>2</sup>Department of Clinical Psychology, University of Amsterdam, Netherlands & Academic Center for Trauma and Personality, Amsterdam, Netherlands | <sup>3</sup>Oxford Centre for Psychological Health, Oxford, UK | <sup>4</sup>Oxford Institute of Clinical Psychology Training and Research, Oxford, UK | <sup>5</sup>Perth Voices Clinic, Murdoch, Western Australia, Australia | <sup>6</sup>School of Psychology, Murdoch University, Murdoch, Western Australia, Australia

**Correspondence:** Georgie Paulik ([georgiepaulik@perthvoicesclinic.com.au](mailto:georgiepaulik@perthvoicesclinic.com.au))

**Received:** 4 July 2024 | **Revised:** 14 August 2024 | **Accepted:** 17 August 2024

**Keywords:** auditory verbal hallucinations | imagery rescripting | post-traumatic stress | PTSD | trauma | voices

## ABSTRACT

**Objective:** People who hear voices (auditory verbal hallucinations) often have post-traumatic stress disorder (PTSD) symptoms. Cognitive behavioural therapies (CBT) have yielded inconsistent findings when treating voices and PTSD symptoms in voice hearers. Preliminary evidence suggests imagery rescripting (ImRs) is associated with large reductions in voice hearing and PTSD symptoms. This study replicated past studies using a larger sample to examine the effectiveness of ImRs in reducing such symptoms.

**Method:** Participants ( $N = 49$ ; 65.3% female;  $M_{age} = 35.86$ ) were clients at an Australian transdiagnostic clinic for voice hearers. A one-arm open trial design was used with three pre-treatment baselines and a mid-treatment, post-treatment and 3-month follow up assessments of PTSD symptoms (Posttraumatic Diagnostic Scale for DSM-5), voices (Hamilton Program for Schizophrenia Voices Questionnaire) and emotional symptoms (Depression Anxiety and Stress Scales-21). Five single-item measures were administered weekly to explore the trajectories of change in trauma intrusions, voice-related distress, voice frequency and positive and negative voice valance.

**Results:** ImRs was associated with very large reductions in PTSD symptoms and voices (both emotional and physical characteristics of voices) and emotional symptoms at post-treatment and follow-up ( $\eta^2_p = 0.24-0.44$ ). There were medium-large to large reductions in weekly symptoms of intrusions, voice-related distress, voice frequency and negative voices ( $\eta^2_p = 0.12-0.16$ ) and a non-significant increase in positive voices ( $\eta^2_p = 0.05$ ).

**Conclusions:** This study provides further evidence that ImRs is an effective treatment for voices and PTSD symptoms in voice hearers with a range of diagnoses. Randomised controlled trials are needed to compare the efficacy of ImRs to CBT protocols.

## 1 | Introduction

‘Voice hearing’ (auditory verbal hallucinations) is characterised by the perception of a voice in the absence of an external stimulus and is the most common type of unusual perceptual experience (Chaudhury 2010). Although voice hearing is typically conceptualised as a positive symptom associated with

psychotic spectrum disorders, voice hearing is experienced in non-psychotic disorders and is relatively common in psychologically healthy individuals (de Leede-Smith and Barkus 2013; Laroï et al. 2012). Voice hearing can therefore be conceptualised as a continuum of symptoms that is distinct from psychotic spectrum diagnoses and other ‘psychotic-like’ symptoms (e.g., delusions; Strachan, Paulik, and McEvoy 2022).

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Author(s). *Clinical Psychology & Psychotherapy* published by John Wiley & Sons Ltd.

### Summary

- Clinicians can target childhood trauma and associated post-traumatic intrusions in psychological therapy to help reduce distressing trauma-related voices.
- Imagery Rescripting may be a safe and effective treatment of post-traumatic stress symptoms in trauma-affected voice hearing individuals.
- Targeting traumatic memories underlying the content of distressing trauma-related voices using Imagery Rescripting may be effective in reducing voice related distress and voice severity.

There is robust evidence for a high prevalence of traumatic life events within people who hear distressing voices (de Bont et al. 2015; Hardy et al. 2016; Tolmeijer et al. 2021). In particular, childhood and adulthood sexual abuse and bullying are strongly associated with voice hearing (Bailey et al. 2018; Hardy et al. 2005). Approximately 75% of trauma-affected voice hearers experience trauma-related voices, which are characterised by voice content that is linked to past trauma (Hardy et al. 2005; Peach et al. 2020). Hardy et al. (2005) reported that voices commonly have a *thematic* link to the past trauma (e.g., voices making malicious remarks to a person with a history of oppression). However, trauma-affected voice hearers often also report voices with a *direct* link to the past trauma, such as voices that repeat the exact comments that were made during an assault (Hardy et al. 2005).

Cognitive behavioural therapies (CBT) that target mechanisms associated with broader positive symptoms (hallucinations, delusions and disorganised cognitions), such as maladaptive beliefs about voices, have shown inconsistent effectiveness in treating voice hearing in psychotic and non-psychotic spectrum populations (Paulik et al. 2019; Thomas et al. 2014). CBT that targets mechanisms of post-traumatic stress disorder (PTSD), including maladaptive trauma appraisals (Zalta 2015), is also ineffective in reducing voice-related distress and is associated with small reductions in PTSD symptoms in samples with psychotic-like symptoms (Brand et al. 2018). However, there is some evidence to suggest that imagery rescripting (ImRs), which is a mental imagery technique that aims to modify the meaning of trauma memories that are linked to current psychological difficulties (Arntz, 2011, 2012), is associated with large reductions in voice frequency, voice-related distress and PTSD symptoms in transdiagnostic voice hearing samples (Ison et al. 2014; Paulik, Steel, and Arntz 2019), with 83% experiencing reliable change and 58.3% experiencing clinically significant change on PTSD symptoms and 25% no longer meeting criteria for PTSD at 1-week follow up (Clarke, Kelly, and Hardy 2022).

Together, these findings suggest that (a) cognitive behavioural models of PTSD and positive symptoms do not adequately explain voice hearing in the context of trauma; (b) there may be differences and similarities in the mechanisms that underlie trauma-related voices, other positive symptoms and PTSD symptoms; and (c) compared to CBT, ImRs may more effectively manipulate mechanisms associated with trauma-related voices. However, there are only three known ImRs treatment studies of trauma-affected voice hearers, each of which have limitations. All three

studies had small sample sizes ( $n \leq 12$ ) and used semi-structured clinician-related interviews to measure voice hearing (Clarke, Kelly, and Hardy 2022; Ison et al. 2014; Paulik et al. 2019), which may introduce clinician bias. Two studies did not assess trauma-related voices specifically (Clarke, Kelly, and Hardy 2022; Ison et al. 2014). Although voice hearing occurs across a range of disorders (de Leede-Smith and Barkus 2013; Larøi et al. 2012), the inclusion criteria of two studies were based on diagnostic status (Clarke, Kelly, and Hardy 2022; Ison et al. 2014), which limits the generalisability of findings to voice hearers who do not meet criteria for psychotic or trauma-spectrum diagnoses. Two studies used a single baseline design without a comparison group (Ison et al. 2014; Paulik, Steel, and Arntz 2019), which limits the ability to confidently conclude that change was attributable to treatment effects, rather than the effects or spontaneous remission, test reactivity or regression to the mean. Two studies used short follow-up periods only (Clarke, Kelly, and Hardy 2022; Ison et al. 2014), which limits the ability to assess delayed effects and the stability of effects. As such, there is a need to extend past studies by testing the impact of ImRs on trauma-related voices within larger transdiagnostic samples and by assessing the stability of symptoms across multiple baselines and the stability of treatment effects over a longer-term follow up period.

Based on past findings, our primary hypotheses were that ImRs would be associated with reductions in the severity of PTSD symptoms and emotional (distressing) characteristics of voices. Our secondary hypotheses were that ImRs would be associated with reductions in PTSD re-experiencing symptoms, physical characteristics of voices (voice frequency/intensity) and emotional symptoms. Additionally, we aimed to explore trajectories of change in symptoms across treatment using weekly measures of the number of trauma-related intrusions, voice frequency, voice-related distress and the frequency of positively and negatively valenced voices.

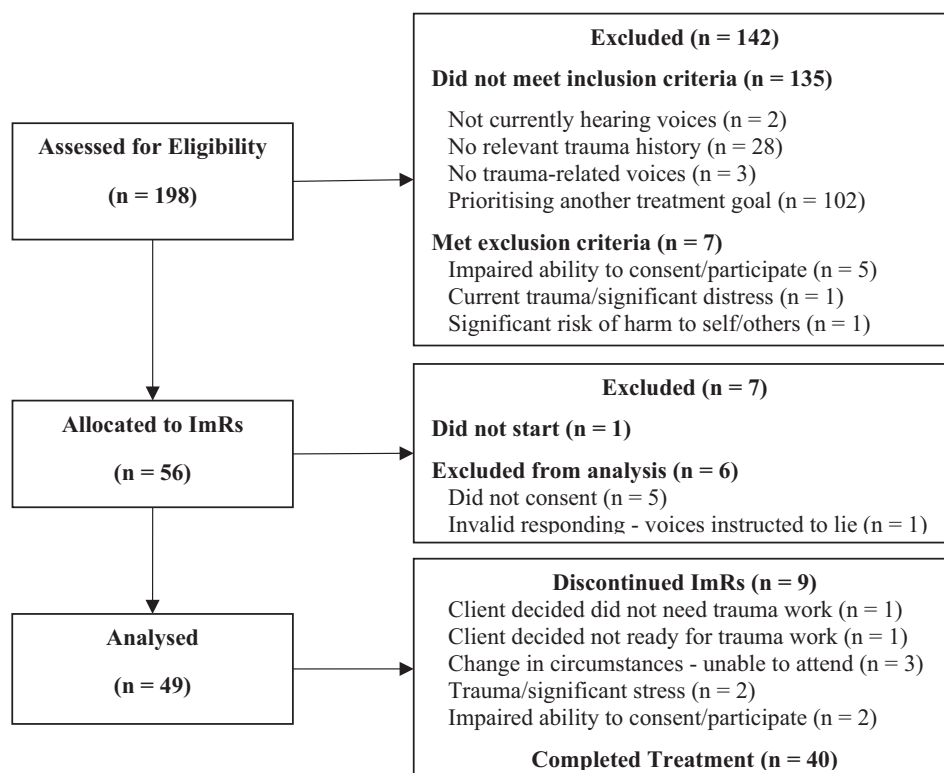
## 2 | Method

### 2.1 | Design

A one-arm open trial design was used with three pre-treatment baseline assessments and one mid-treatment, post-treatment and 3-month follow up assessments of full psychometric measures. A short battery of single-item measures was administered at each of the three weekly baseline assessments, immediately prior to each treatment session and at 3-month follow up.

### 2.2 | Participants

Participants ( $N=49$ ) were help-seeking clients at an Australian specialist clinic for people who hear voices (Perth Voices Clinic), who received ImRs as part of routine service provision. Inclusion criteria for ImRs are: (a) currently (within the past two weeks) hearing voices that are causing distress, (b) a trauma history associated with current post-traumatic stress symptoms,<sup>1</sup> (c) voices that are indirectly or directly linked to trauma (clinician-assessed), (d) client and clinician agreement that trauma symptoms are the primary treatment goal and (e)  $\geq 18$  years of age. Exclusion criteria are: (a) experiencing acute psychosis, enduring



**FIGURE 1** | CONSORT flow diagram for ImRs.

delusions, thought disorder (scored  $\geq 5$  on unusual thought content or conceptual disorganisation items of the Brief Psychiatric Rating Scale; Ventura et al. 1993) or another psychological condition that the clinician believed may impair participants' ability to consent or engage effectively in treatment (e.g., dementia), (b) in a current mental health or situational crisis (especially experiencing trauma that was related to the trauma being targeted in ImRs) at the time of therapy commencement or (c) at significant risk of harming themselves or others. ImRs clients who did not consent for their data to be used for service evaluation were excluded from analyses (see Figure 1 for CONSORT flow diagram). Demographic and clinical characteristics are reported in Table 1.

## 2.3 | Measures

### 2.3.1 | Primary Outcomes

**2.3.1.1 | PTSD Symptoms.** The Posttraumatic Diagnostic Scale for DSM-5 (PDS-5) comprises 22-items with a five-point scale ranging from 0 (*not at all*) to 4 (*six or more times a week/severe*; Foa et al. 2016). The PDS-5 total score was used to measure the severity of PTSD symptoms.

**2.3.1.2 | Emotional Characteristics of Voices.** The emotional (distress) subscale of the Hamilton Program for Schizophrenia Voices Questionnaire (HPSVQ-E) was used to measure emotional characteristics of voices (Kim et al. 2010; Van Lieshout and Goldberg 2007). The HPSVQ is a 13-item measure with a five-point rating scale ranging from 0 (*least severe/impairing*) to 4 (*most severe/causes the largest amount of disruption*). Items 10–13 are not included in the distress scores and were not administered.

### 2.3.2 | Secondary Outcomes

**2.3.2.1 | PTSD Re-experiencing Symptoms.** The re-experiencing subscale of the PDS-5 (PDS-5-RE) was used to measure the severity of PTSD re-experiencing symptoms (i.e. intrusive cognitions, emotions, sensations).

**2.3.2.2 | Physical Characteristics of Voices.** The physical (frequency/intensity) subscale of the HPSVQ (HPSVQ-P) was used to measure physical characteristics of voices (Kim et al. 2010; Van Lieshout and Goldberg 2007).

## 2.4 | Emotional Symptoms

The Depression Anxiety and Stress Scales (DASS-21) is a 21-item measure that provides a four-point rating scale ranging from 0 (did not apply to me at all) to 3 (applied to me very much, or most of the time; Lovibond and Lovibond 1995). The DASS-21 total score was used to measure overall emotional symptoms.

### 2.4.1 | Weekly Symptoms

A short battery of five single-item questions was developed by authors to assess trauma-related symptoms on a weekly basis. The 'intrusions' item, 'how many trauma-related intrusions (i.e., nightmares, flashbacks) did you experience in the past week?', asked participants to estimate the total number of PTSD re-experiencing symptoms. The 'voice frequency' item, 'in the past week, including today, how frequently did you experience voices?', measured voice frequency on a 7-point rating scale ranging from 0 (*voices not present or present less than once a week*) to 6

**TABLE 1** | Participant demographic and clinical characteristics ( $N = 49$ ).

	<b><i>N</i> (%) participants</b>
Age (years)	$M = 35.86$ ( $SD = 14.33$ , range = 18–67)
Gender	
Female	32 (65.3%)
Male	13 (26.5%)
Non-binary	4 (8.2%)
Employment status	
Employed paid full-time	11 (22.5%)
Unemployed	32 (65.3%)
Student	6 (12.2%)
Relationship status	
Single	31 (63.3%)
Married/long-term partner	18 (36.7%)
Education	
Left school $\leq 16$ years old	16 (32.7%)
Completed/completing Year 12	5 (10.2%)
Completed/completing college or university	28 (57.1%)
Currently taking psychotropic medications	42 (85.7%)
Age at voice hearing onset	20.17 ( $SD = 12.72$ , range = 5–52)
Previously received psychological treatment	43 (87.8%)
Referrers' diagnostic impressions <sup>a</sup>	
Psychotic spectrum disorder	22 (44.9%)
Non-psychotic diagnosis	27 (55.1%)
Other mental health condition than psychosis	16 (32.7%)
Comorbidity	34 (69.4%)
Types of traumas <sup>b</sup>	
Serious, life-threatening illness	4 (8.2%)
Physical assault	22 (44.9%)
Sexual assault	15 (30.6%)
Child non-sexual abuse	37 (75.5%)
Child sexual abuse	24 (49%)
Accident	6 (12.2%)
Other trauma	33 (67.3%)
Number of different types of traumas	2.88 ( $SD = 1.24$ )

(Continues)

**TABLE 1** | (Continued)

	<b><i>N</i> (%) participants</b>
Trauma–voice association	
Indirect only	21 (42.9%)
Direct only	2 (4.1%)
Both	26 (53.0%)

Abbreviations:  $M$  = mean,  $SD$  = standard deviation.<sup>a</sup>Diagnostic information is from referral letters and may not have been made as part of a formal diagnostic assessment.<sup>b</sup>As reported by client on the PDS-5 Part A.

(*voices occur continuously or almost continually*). The 'voice distress' item, 'on average, over the past week (including today), how distressed did you get by your voices?', measured voice-related distress using a rating scale that ranged from 0 (*no distress*) to 100 (*maximum distress*). The 'positive valence' and 'negative valence' items, 'in the past week, how often were your voices positive?' and 'in the past week, how often were your voices negative?', measured positive and negative voices content using a 5-point response scale that ranged from 0 (*never*) to 4 (*a lot*).

## 2.5 | Procedure

The [Murdoch University] Human Research Ethics Committee approvals were obtained prior to data collection (2016/089; HRE2019-0525). The number of sessions were based on Medicare's (Australian national insurance scheme) funding model for private clinical psychology services. At the start of the trial, Medicare funded 10 sessions per year, which were used for one assessment session, one psychoeducation and ImRs preparation session, seven weekly rescripting sessions (i.e., of one memory) and one wrap-up session. In August 2020, Medicare-funded sessions were increased to 20 per year in response to the impact of COVID-19, and participants who commenced treatment after this time received between seven to 17 ImRs sessions ( $M = 8.06$ ;  $Median = 7$ ), depending on their clinical needs. Participants (36.7%) most frequently received seven sessions, with 20.4% receiving less than seven, 20.4% receiving 8–10, 10.2% receiving 11–13 and 12.2% receiving 14–17 sessions. Sessions were delivered via telehealth or face-to-face.

### 2.5.1 | Session Content

Session 1 (assessment) involved assessment of current and past mental health history (primarily voice hearing and other psychosis-related experiences), trauma history, inclusion and exclusion criteria and trauma–voice associations (Hardy et al. 2005). Session 2 (ImRs preparation) occurred 2 weeks after session one and involved psychoeducation, sequencing of selected trauma memories for rescripting, and a homework visualisation task of imagining themselves soothing their former child self. Sessions 3–19 (rescripting) occurred weekly and involved rescripting of one memory per session. In the rescripting protocol (detailed protocol available on request), initially the participant brings a trauma memory to mind and describes in present tense, from the perspective of their former child self, what is happening in the image, their sensory, cognitive and emotional experiences

at the start of the memory. Once the client is connected to the feelings evoked by the memory, but typically before the 'hot' part of the trauma has taken place, the therapist intervenes in the event and supports the participant to meet their needs. The therapist describes how they are intervening and instructs the participant to imagine them carrying out these actions. The therapist provides comfort, reassurance, explicitly tells the child they are not guilty and that the offender should be ashamed. The therapist continues to intervene until the participant feels safe. Due to the complexity of trauma amongst voice hearers and increased likelihood of dissociation in later stages of ImRs, this trial did not include the second and third stages of ImRs that involve participants (adult self) intervening in the imagined event (Arntz and Weertman 1999). The final post-treatment (wrap-up) session occurred 1 week after the final rescripting session and involved treatment review, relapse prevention, and discussion of future goals. All interventions were delivered by G. P. (author).

### 2.5.2 | Assessment Time Points

Assessment of full-scale measures occurred at six time points: baseline one (BL1; session one), baseline two (BL2; between session one and two), baseline three (BL3; session two), mid-treatment (MT; before Session 6; based on initial Medicare-funded treatment plan), post-treatment (PT) and 3-month follow-up (FU). Consent for service evaluation data to be used in research was requested at baseline two. To limit participant burden, the secondary outcome (DASS-21) was excluded from baselines two and three. Assessment of single items occurred at all baseline assessments, all active rescripting sessions (IR1–IR17), mid-treatment, post-treatment and follow-up. Data collection occurred between September 2019 and December 2022 using Qualtrics. Participants were provided a link to complete measures online directly via Qualtrics or given paper copies, which were uploaded to Qualtrics by the research team. All measures were completed after treatment sessions.

## 2.6 | Data Analysis

A statistical analysis plan was uploaded to Open Science Framework during the final stages of data collection ([https://osf.io/pqyrt/?view\\_only=ee2cdd0e86f3440490711ef404dd492a](https://osf.io/pqyrt/?view_only=ee2cdd0e86f3440490711ef404dd492a); view only link for blind review). A G\*Power (Faul et al. 2007) a priori power analyses, assuming medium effects ( $f=0.25$ ), 80% power, six, three and 22 measurement occasions for primary, secondary and exploratory outcomes, respectively, and alphas of 0.05 indicated that a minimum of 28 participants were needed.

### 2.6.1 | Treatment Efficacy

A series of linear mixed models (LMMs) using data from the full intention-to-treat sample were used to test the impact of ImRs on primary and secondary outcomes. Full information estimation used all available data at each timepoint to account for missing data. Time was entered as categorical fixed factor and the number of active rescripting sessions was entered as a fixed covariate. To account for individual differences in baseline scores, a random intercept was included for each participant. Time was

entered as a random slope but was removed from the models if it caused convergence issues. The fit of models with (AR1) and without (unstructured) autoregressive residual covariance structure was compared using BIC statistic. Covariance structures with the best fit were retained. Partial eta-square was used to index the effect size for the main effect of time. Cohen's  $d$  (denominator = pooled SD) was used to index effect sizes for least significant difference contrast coefficients, which were used to test differences between model-estimated marginal means across assessments, including assessment of baseline stability.

**2.6.1.1 | Weekly Measures.** Three phases were distinguished: baseline, ImRs and 3-month follow-up. As weekly measures were taken before the session and asked for the last week, assessments 1–4 were taken as baseline, and assessment 5 was post-treatment as ImRs phase. Six participants that discontinued before active ImRs started were not included in the analyses of the weekly measures, leaving  $N=43$ . Within each phase, time was centred on an individual basis. Data were analysed with multilevel analysis, using SPSS29 (Generalised) Linear Mixed Models ((G)LMM), depending on which had superior fit, an AR1 or an ARMA11 covariance structure for the repeated part. If estimation allowed (i.e., if convergence was reached) a random intercept for participant was added. The fixed part consisted of intercept, phase (with baseline as reference) as factor, and centred time-within-baseline and time-within-ImRs as dimensional predictors. In case a within phase-centred time predictor was nonsignificant at  $p=0.10$ , it was deleted from the model. Based on type of variable (i.e., counts or ratings) and distribution, an appropriate distribution was chosen (negative binomial with loglink for counts and normal for ratings). In case of negatively skewed data, scores were inverted to enable analysis with negative binomial regression.

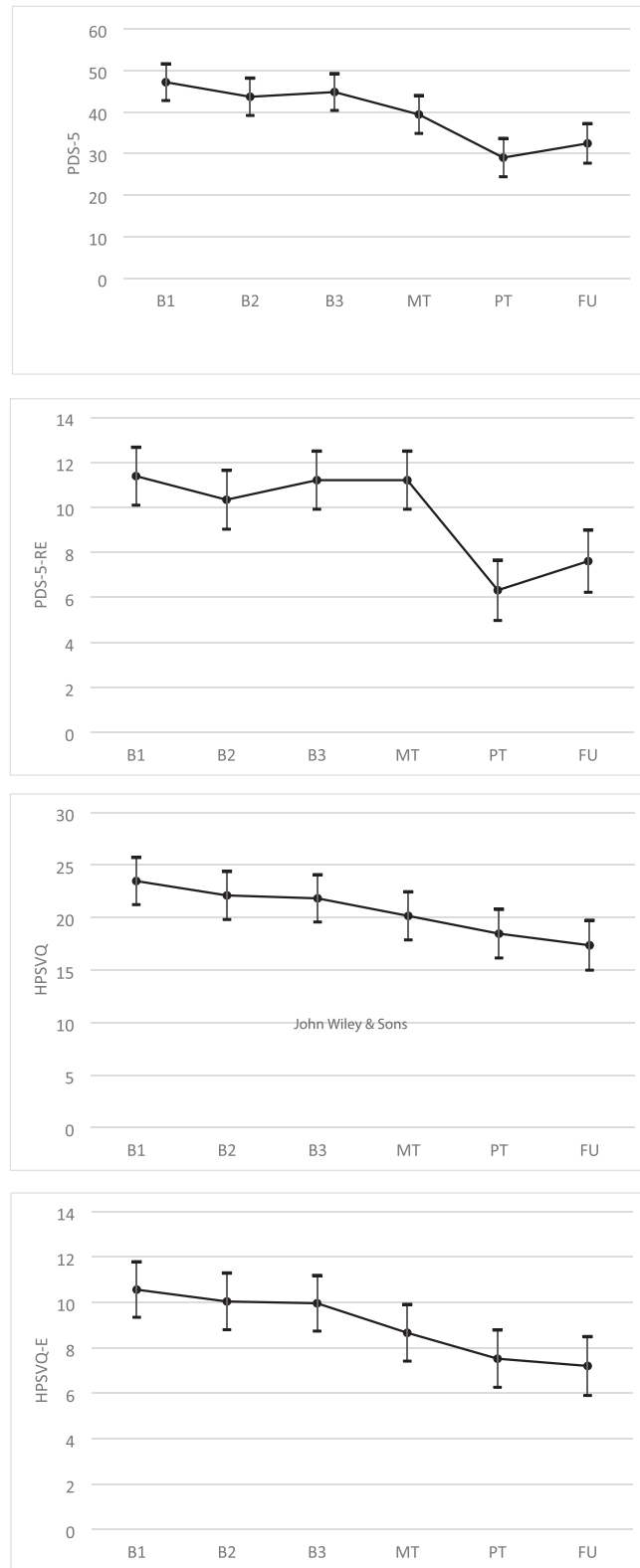
### 2.6.2 | Reliable and Clinically Significant Change

To evaluate whether individual change in PTSD and voice hearing severity from B1 to MT, PT and FU were statistically significantly higher than chance due to measurement error, Jacobson and Truax's (1992) reliable change index (RCI) scores were calculated for the PDS-5 and HPSVQ (total score). Jacobson and Truax's (1992) clinically significant change (CSC) statistic was used to evaluate clinically meaningful change in voice hearing severity from B1 to MT, PT and FU. The threshold for clinically significant change was two standard deviations ( $SD=6.00$ ) below the clinical mean ( $M=18.80$ ; Van Lieshout and Goldberg 2007). As no disaggregated clinical versus non-clinical norms are available for the PDS-5, change in the probable clinical cut-off for PTSD diagnosis (PDS-5 total score  $\geq 28$ ; Foa et al. 2016) was calculated as an alternative index of clinical change.

## 3 | Results

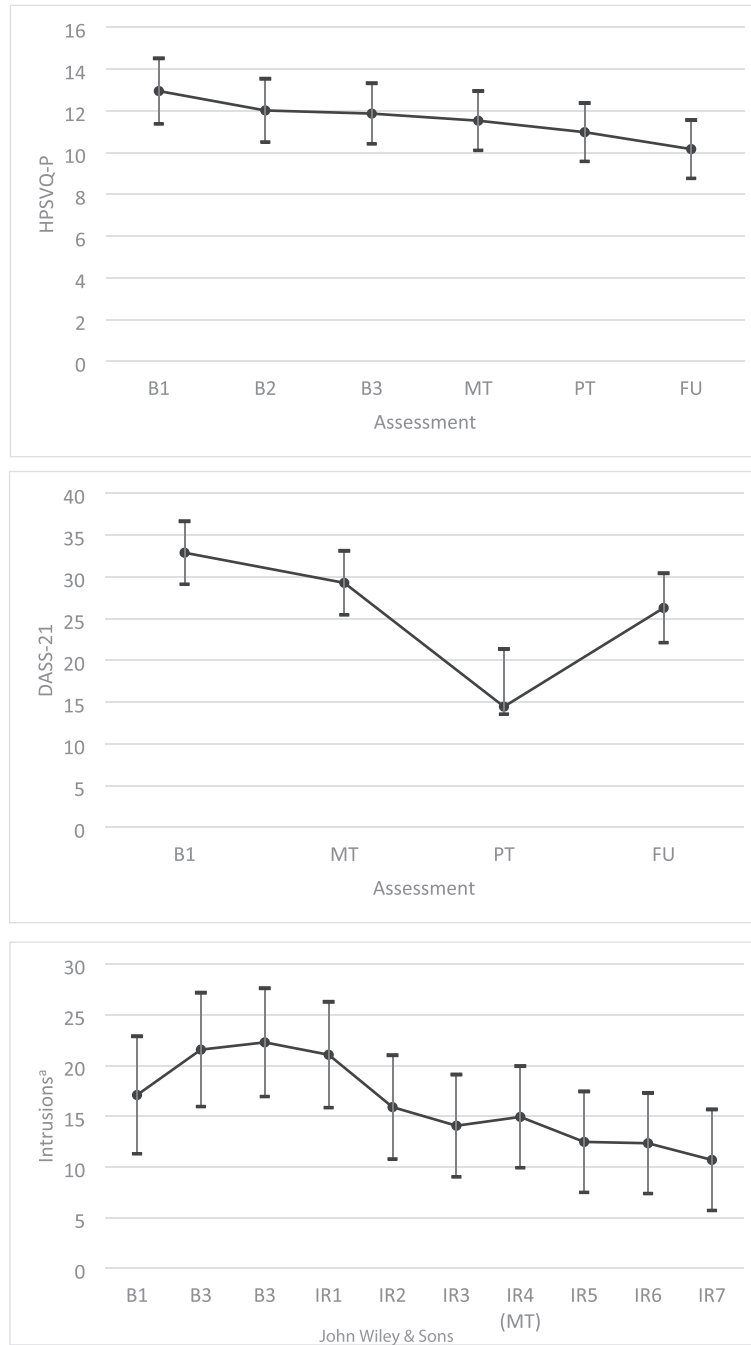
### 3.1 | Treatment Outcomes Across Time

Three weekly measures were skewed so negative binomial regression with loglink was used. All other models met normality assumptions. A random slope for time was retained in the physical characteristics of voices (HPSVQ-P) but was removed



**FIGURE 2** | Plots of model-predicted means for primary, secondary, weekly and mechanism outcomes. Abbreviations: B=baseline, DASS=Depression Anxiety and Stress Scales-21, FU=3-month follow-up, HPSVQ=Hamilton Program for Schizophrenia Voices Questionnaire, HPSVQ-E=Hamilton Program for Schizophrenia Voices Questionnaire - Emotional Characteristics subscale, HPSVQ-P=Hamilton Program for Schizophrenia Voices Questionnaire - Physical Characteristics subscale, Intrusions=weekly intrusion frequency, IR=active imagery rescripting session, MT=mid-treatment, Negative valence=weekly frequency of negatively valenced voices, PDS-5=Posttraumatic Diagnostic Scale for DSM-5, PDS-5-RE=Posttraumatic Diagnostic Scale for DSM-5 - Re-experiencing subscale, Positive valence=weekly frequency of positively valenced voices; PT=post-treatment; Voice distress=weekly voice-related distress; Voice frequency=weekly frequency of voices. <sup>a</sup>Plots of weekly measures do not include ImRs sessions 8–17, given that few participants completed later ImRs sessions and plotted data from these sessions may provide a biased representation of the larger sample.

FIGURE 2 | (Continued)



from the remaining models due to convergence problems. The number of ImRs sessions was a non-significant covariate in all models ( $ps \geq 0.06$ ). LMM outputs, raw means and standard deviations, Cronbach's alphas, bivariate correlations between all variables and model-estimated means are available at [https://osf.io/pqyrt/?view\\_only=ee2cdd0e86f3440490711ef404dd492a](https://osf.io/pqyrt/?view_only=ee2cdd0e86f3440490711ef404dd492a) (view only link for blind review). Model-estimated means for each assessment point are plotted in Figure 2.

### 3.1.1 | Baseline Stability

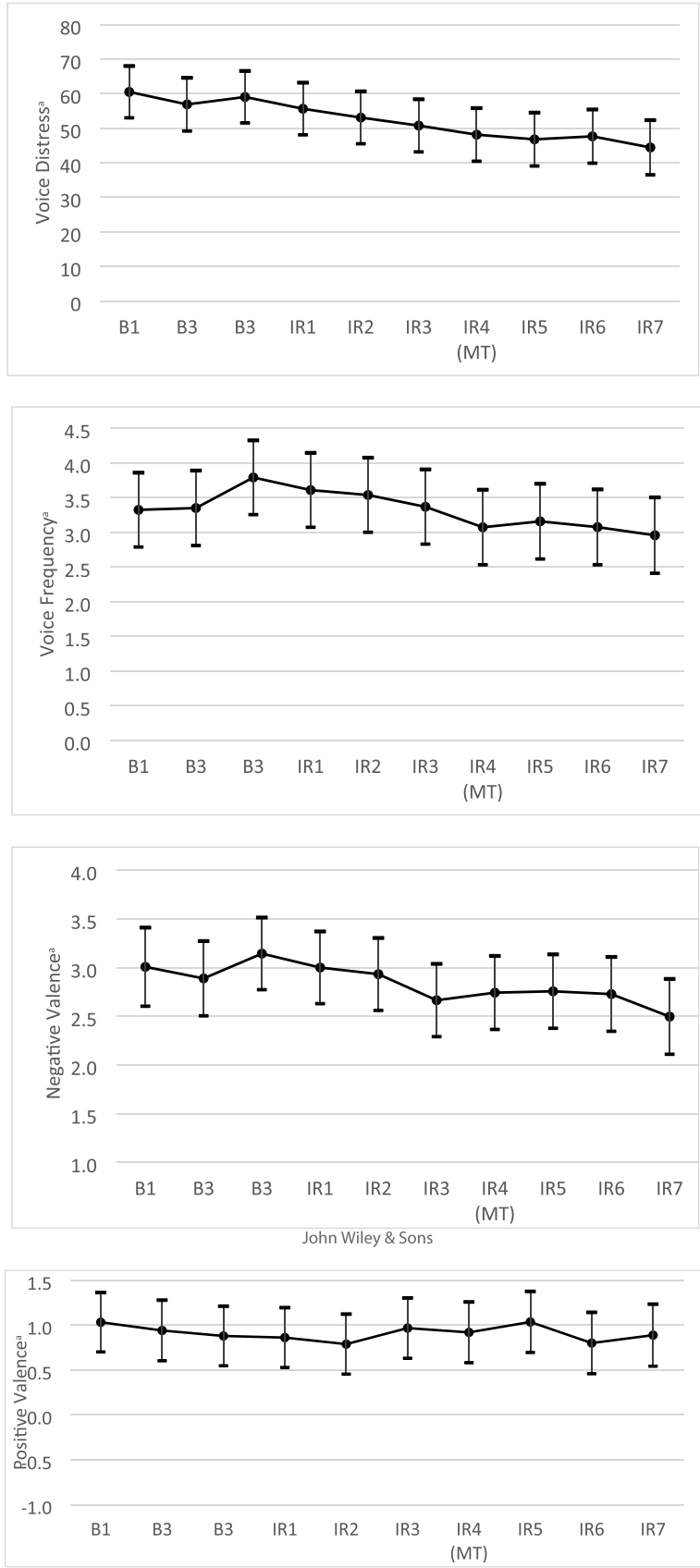
Across the three baseline assessments, there were two significant changes in the physical characteristics of voices and

weekly intrusions and one in weekly voice frequency, which were small ( $ds = 0.02-0.27$ ). There were no other significant baseline differences, which suggests that outcomes were relatively stable prior to active ImRs sessions. Mean change and effect sizes across baseline assessments are presented in Table 2.

### 3.1.2 | Primary Outcomes

There were significant and very large main effects of time on the severity of PTSD symptoms and emotional characteristics of voices: PDS,  $F(5, 205.89) = 24.11$ ,  $p < 0.001$ ,  $\eta^2_p = 0.37$ , and HPSVQ-E,  $F(5, 205.57) = 15.01$ ,  $p < 0.001$ ,  $\eta^2_p = 0.27$ . There was significant and very large main effect

FIGURE 2 | (Continued)



**TABLE 2** | Model-estimated mean changes with effect sizes across baselines.

	Mean change from BL1						Mean change from BL2		
	BL2			BL3			BL3		
	$M_{ch}$	$p$	$d$	$M_{ch}$	$p$	$d$	$M_{ch}$	$p$	$d$
PDS-5	-3.51 (1.93)	0.07	0.22	-2.37 (1.89)	0.21	0.15	1.14 (1.95)	0.56	0.07
PDS-5-RE	-1.05 (0.55)	0.06	0.22	-0.18 (0.34)	0.74	0.03	0.87 (0.55)	0.06	0.19
HPSVQ	-1.38 (0.86)	0.11	0.17	-1.66 (0.83)	0.05	0.21	-0.28 (0.86)	0.75	0.04
HPSVQ-E	-0.52 (0.78)	0.28	0.12	-0.60 (0.46)	0.19	0.14	-0.08 (0.48)	0.86	0.02
HPSVQ-P	-0.92 (0.36)	0.01	0.17	-1.07 (0.42)	0.01	0.20	-0.15 (0.36)	0.67	0.03
Intrusions	4.47 (2.06)	0.03	0.10	5.19 (2.15)	0.02	0.22	0.71 (2.02)	0.72	0.13
Voice distress	-3.62 (3.50)	0.30	0.07	-1.46 (3.40)	0.67	0.04	2.16 (3.50)	0.54	0.12
Voice frequency	0.03 (0.17)	0.88	0.09	0.47 (0.17)	0.006	0.31	0.44 (0.18)	0.01	0.21
Positive valence	-0.09 (0.15)	0.54	0.06	-0.15 (0.15)	0.30	0.12	0.06 (0.15)	0.69	0.06
Negative valence	-0.12 (0.16)	0.47	0.06	0.14 (0.18)	0.44	0.25	0.26 (0.17)	0.12	0.29

Abbreviations: BL = baseline, HPSVQ = Hamilton Program for Schizophrenia Voices Questionnaire, HPSVQ-E = Hamilton Program for Schizophrenia Voices Questionnaire - Emotional Characteristics subscale, HPSVQ-P = Hamilton Program for Schizophrenia Voices Questionnaire - Physical Characteristics subscale, Intrusions = weekly intrusion frequency,  $M_{ch}$  = estimated mean change, Negative valence = weekly frequency of negatively valenced voices, PDS-5 = Posttraumatic Diagnostic Scale for DSM-5, PDS-5-RE = Posttraumatic Diagnostic Scale for DSM-5, Positive valence = weekly frequency of positively valenced voices, Voice distress = weekly voice-related distress, Voice frequency = weekly frequency of voices.

on the overall severity of voices (total score), HPSVQ,  $F(5, 204.61) = 12.66$ ,  $p < 0.001$ ,  $\eta^2_p = 0.24$ . Both primary outcomes showed significant medium to very large reductions at post-treatment and follow-up ( $ds = 0.69-1.14$ ). There were no significant changes in either primary outcome from post-treatment to follow-up.

### 3.1.3 | Secondary Outcomes

There were significant and very large main effects of time on the severity of PTSD re-experiencing symptoms, the physical characteristics of voices and emotional symptoms: PDS-5-RE,  $F(5, 211.01) = 26.87$ ,  $p < 0.001$ ,  $\eta^2_p = 0.39$ , HPSVQ-P,  $F(5, 205.57) = 15.01$ ,  $p < 0.001$ ,  $\eta^2_p = 0.27$ , and DASS-21,  $F(3, 112.73) = 29.65$ ,  $p < 0.001$ ,  $\eta^2_p = 0.44$ . There were significant medium to very large reductions ( $ds = 0.38-2.23$ ) on secondary outcomes at post-treatment and follow-up. Although there was a further significant and small reduction in the physical characteristics of voices from post-treatment to follow-up ( $d = 0.17$ ), there were significant large and medium-large increases in PTSD re-experiencing symptoms ( $d = 1.18$ ) and emotional symptoms ( $d = 0.62$ ), respectively. Mean change and effect sizes between baseline one, mid- and post-treatment and follow-up for all variables are reported in Table 3.

### 3.1.4 | Weekly Measures

Table 4 presents the results of the multilevel analyses. The frequency of 'negative valence' voices was negatively skewed; hence, ratings were inverted before analysed. For three variables, the slope during baseline was not significant, hence deleted from the model. For 'voice frequency', frequency of 'positive valence' voices deteriorated during baseline at  $p$  levels between 0.05 and 0.10; hence these effects were maintained in the model. All but one (the frequency of 'positive valence' voices) weekly measures showed significant improvement halfway through ImRs as well as at 3-month follow-up. Slopes during ImRs all indicated significant improvement. In general, at 3-month follow-up, effects were smaller than at end of ImRs, yet still significant.

### 3.1.5 | Reliable and Clinically Significant Change

Table 5 presents the proportion of participants who showed reliable and clinically significant change on the PDS-5 and HPSVQ at mid-treatment, post-treatment and follow-up. There were high rates of reliable change on the PDS-5 ( $> 49\%$ ), low rates of reliable deterioration ( $< 8\%$ ) and the proportion of participants who met the cut-off for probable PTSD more

**TABLE 3 |** Model-estimated mean changes with effect sizes for all variables across treatment and 3-month follow-up assessments.

	Mean change from BL1						Mean change from MT						Mean change from PT					
	MT			PT			FU			PT			FU			FU		
	$M_{ch}$	$p$	$d$	$M_{ch}$	$p$	$d$	$M_{ch}$	$p$	$d$	$M_{ch}$	$p$	$d$	$M_{ch}$	$p$	$d$	$M_{ch}$	$p$	$d$
PDS-5	-7.77 (1.98)	<0.001	0.49	-18.15 (2.01)	<0.001	1.14	-14.73 (2.10)	<0.001	0.89	-10.38 (2.06)	<0.001	0.64	-6.96 (2.15)	0.001	0.42	3.42 (2.17)	0.12	0.21
PDS-5-RE	-0.18 (0.54)	0.74	0.03	-5.07 (0.57)	<0.001	1.08	-3.78 (0.59)	<0.001	2.23	-4.89 (0.57)	<0.001	1.05	-3.60 (0.59)	<0.001	2.24	1.30 (0.61)	0.04	1.18
HPSVQ	-3.32 (0.86)	<0.001	0.42	-5.01 (0.88)	<0.001	0.62	-6.12 (0.92)	<0.001	0.74	-1.69 (0.89)	0.06	0.21	-2.80 (0.93)	0.003	0.34	-1.10 (0.94)	0.24	0.13
HPSVQ-E	-1.90 (0.48)	<0.001	0.44	-3.04 (0.49)	<0.001	0.69	3.37 (0.51)	<0.001	0.74	-1.14 (0.50)	0.02	0.26	-1.47 (0.52)	0.005	0.33	-0.33 (0.53)	0.54	0.07
HPSVQ-P	-1.42 (0.48)	0.003	0.27	-1.97 (0.52)	<0.001	0.38	-2.78 (0.57)	<0.001	0.57	-0.55 (0.38)	0.15	0.11	-1.36 (0.47)	0.004	0.28	-0.81 (0.40)	0.04	0.17
DASS-21	-3.61 (1.66)	0.03	0.28	-15.43 (1.71)	<0.001	1.15	-6.61 (1.86)	<0.001	0.47	-11.82 (1.69)	<0.001	0.87	-3.00 (1.84)	0.11	0.21	8.82 (1.86)	<0.001	0.62
Intrusions	-2.16 (2.52)	0.39	0.09	-11.03 (2.76)	<0.001	0.59	-4.69 (2.87)	0.10	0.19	-8.88 (2.75)	0.001	0.57	-2.53 (2.86)	0.38	0.10	6.34 (2.26)	0.005	0.62
Voice distress	-12.37 (3.52)	<0.001	0.37	-22.12 (3.60)	<0.001	0.77	-17.31 (3.77)	<0.001	0.51	-9.74 (3.68)	0.008	0.37	-4.94 (3.85)	0.20	0.17	4.80 (3.91)	0.22	0.17
Voice frequency	0.25 (0.18)	0.16	0.09	-0.70 (0.18)	<0.001	0.37	-0.73 (0.19)	<0.001	0.37	-0.44 (0.18)	0.02	0.30	-0.48 (0.19)	0.01	0.30	-0.03 (0.20)	0.87	0.02
Positive valence	-0.11 (0.15)	0.46	0.08	0.14 (0.16)	0.36	0.12	0.12 (0.16)	0.46	0.11	0.25 (0.16)	0.11	0.20	0.23 (0.17)	0.16	0.19	-0.02 (0.17)	0.90	0.03
Negative valence	-0.27 (0.19)	0.16	0.13	-0.76 (0.19)	<0.001	0.53	-0.82 (0.20)	<0.001	0.60	-0.42 (0.19)	0.03	0.36	-0.48 (0.20)	0.02	0.42	-0.06 (0.18)	0.75	0.06

Abbreviations: BL = baseline, DASS = Depression Anxiety and Stress Scales-21, FU = 3-month follow-up, HPSVQ-E = Hamilton Program for Schizophrenia Voices Questionnaire – Emotional Characteristics subscale, HPSVQ-P = Hamilton Program for Schizophrenia Voices Questionnaire – Physical Characteristics subscale, Intrusions = weekly intrusion frequency,  $M_{ch}$  = estimated mean change, MT = mid-treatment, PDS-5 = Posttraumatic Diagnostic Scale for DSM-5, PDS-5-RE = Posttraumatic Diagnostic Scale for DSM-5 – Re-experiencing subscale, HPSVQ = Hamilton Program for Schizophrenia Voices Questionnaire, Positive valence = weekly frequency of positively valenced voices, Negative valence = weekly frequency of negatively valenced voices, PT = post-treatment, Voice distress = weekly voice-related distress, Voice frequency = weekly frequency of voices.

TABLE 4 | Results of multilevel analysis of weekly measure.

Distribution	Covariance structure repeated part	Random intercept	Intercept (i.e., baseline mean)			ImRs (mid-treatment)			3-month follow-up			Slope within baseline	
			B	t	p	B	t	p	B	t	p	B	B
Negative binomial (with loglink)	ARMA11	Yes	2.824	18.40	<0.001	-0.819	-7.77	<0.001	-0.623	-3.04	0.003	N.A.	N.A.
Normal	ARMA11	Yes	57.13	16.57	<0.001	-11.59	6.06	<0.001	-14.76	-4.36	<0.001	N.A.	N.A.
Normal	ARMA11	Yes	3.547	13.12	<0.001	-0.389	-3.71	<0.001	-0.894	-4.24	<0.001	0.109	0.109
Negative binomial (with loglink)	ARMA11	No	-0.089	-0.578	0.57	0.039	0.42	0.68	0.205	1.09	0.28	-0.09	-0.09
Negative binomial (with loglink)	ARMA11	No	-0.017	-0.124	0.90	0.318	3.88	<0.001	0.616	4.07	<0.001	N.A.	N.A.

Note: Covariance structure of the repeated part was determined on best fit of AR1 vs. ARMA11. If estimation allowed, a random intercept for participant was added. Time is expressed in weeks. If the slope of time within baseline was not significant at the  $p=0.10$  level, it was deleted from the fixed part (indicated by N.A. = not applicable). Main effects of Intercept (i.e., the reference category, baseline) and ImRs represent the estimated means halfway that phase, that is, at centred within-phase time = 0. Effects of analyses based on negative binomial distribution with loglink are in transformed scale. The weekly negative voices frequency was inverted to enable analysis with negative binomial distribution. This means that the N.S. effect of the intercept indicated that there was no significant difference during baseline with the maximum frequency score (4) and that the positive B's indicate reductions on the original frequency scale.

than halved (<39%). There were lower rates of reliable improvement (>20%) and deterioration (2%) on the HPSVQ, and 24% and 16% of participants experienced clinically significant change at post-treatment and follow-up, respectively.

#### 4 | Discussion

This study examined the impact of ImRs for trauma memories on the severity of emotional characteristics of voices, PTSD symptoms and other trauma-related voice characteristics. The findings from this study support the proposed hypothesis, providing further evidence that ImRs effectively reduces emotional characteristics of voices and PTSD symptoms amongst trauma-affected voice hearers (Clarke, Kelly, and Hardy 2022; Ison et al. 2014; Paulik, Steel, and Arntz 2019).

The multiple baseline design of the current study indicated that observed changes during the treatment phase were likely the result of the active intervention rather than the expected effects of change over time (Hawkins et al. 2007). Also, outcomes from the 3-month follow-up assessment demonstrated significant post-treatment effects (Llewellyn-Bennett, Bowman, and Bulbulia 2016). The current transdiagnostic sample was larger than that of previous voice-hearer ImRs studies ( $ns < 12$ ) and provided evidence that ImRs effectively reduces PTSD symptom severity and the emotional characteristics of voices across a range of disorders.

ImRs was associated with medium-large reductions in the physical characteristics of voices, with a further small reduction from post-treatment to 3-month follow-up, which suggests that there are continued reductions in the loudness, duration, frequency and clarity of voices in the months after ImRs is ceased. The medium-large to large reductions in the severity of voices, the emotional characteristics of voices and voice-related distress were comparable to the findings of past ImRs studies (Paulik, Steel, and Arntz 2019), as was the rate of reliable change in the severity of voices (Ison et al. 2014). The observed effect sizes are somewhat smaller than those reported by Paulik, Steel, and Arntz (2019), which may be due to the use of different voice hearing measures. Consistent with past findings (Paulik, Steel, and Arntz 2019), there were large reductions in the severity of total PTSD symptoms at post-treatment and 3-month follow-up and medium to large reductions in emotional symptoms. Although the rate of reliable change in PTSD symptoms at post-treatment was somewhat lower than that in past ImRs studies with trauma-affected individuals with psychotic spectrum disorders (57% vs. 83.3%; Clarke, Kelly, and Hardy 2022), the proportion of participants who fell below the cut-off for PTSD at post-treatment was comparable across studies. Furthermore, 65% of participants fell below the cut-off for PTSD at 3-month follow-up, which is more than double the proportion who no longer met criteria for PTSD at 1-month follow-up in past ImRs studies in sample that met diagnostic criteria for PTSD and a psychotic spectrum disorder (25%; Clarke, Kelly, and Hardy 2022). However, another study used a different self-report measure of PTSD to compare the effectiveness of eye movement desensitisation and reprocessing (EMDR) and prolonged exposure (PE) in a sample that met diagnostic criteria for PTSD and a psychotic

**TABLE 5** | Proportion (percentage) of total sample ( $N=49$ ) with reliable change, clinically significant change and probable post-traumatic stress disorder at mid-treatment, post-treatment and 3-month follow-up.

	BL1	MT	PT	FU
HPSVQ	—	$n = 42$	$n = 39$	$n = 34$
Reliably improved	—	11 (22%)	12 (24%)	10 (20%)
Reliably deteriorated	—	0 (0%)	0 (0%)	1 (2%)
Clinically significant change	—	0 (0%)	12 (24%)	8 (16%)
PDS	—	$n = 42$	$n = 40$	$n = 35$
Reliably improved	—	17 (35%)	28 (57%)	24 (49%)
Reliably deteriorated	—	3 (6%)	2 (4%)	4 (8%)
Probable PTSD <sup>a</sup>	$n = 40$	$n = 39$	$n = 40$	$n = 35$
	36 (73%)	29 (59%)	19 (39%)	17 (35%)

Abbreviations: BL = baseline, FU = 3-month follow-up, HPSVQ = Hamilton Program for Schizophrenia Voices Questionnaire, MT = mid-treatment,  $n$  = number of participants included in calculation, PDS-5 = Posttraumatic Diagnostic Scale for DSM-5, PT = post-treatment.

<sup>a</sup>As normative data for the PDS-5 in disaggregated clinical vs. non-clinical populations has not been published, clinically significant change was unable to be calculated for the PDS. Change in clinical cut-off for probable PTSD (PDS-5 total score  $\geq 28$ ) was calculated as an alternative index of clinical change.

spectrum disorder. This study reported 78% (EMDR) and 74% (PE) of participants fell below the cut-off for PTSD at 6-month follow-up (van den Berg et al. 2015). Although the differences between and within studies could be due to differences in assessment tools or clinical characteristics, it suggests that other PTSD treatments may be more effective in reducing PTSD symptoms within individuals that meet diagnostic criteria for PTSD and a psychotic spectrum disorder. Future studies are needed to compare the effectiveness of ImRs to other PTSD treatments in reducing PTSD symptoms in transdiagnostic and clinical samples.

In relation to PTSD re-experiencing symptoms specifically, at post-treatment, there were medium-large reductions on the weekly frequency of trauma-related intrusions single-item, which is consistent with past findings (Paulik et al. 2019) and large reductions on the validated full-scale measure of PTSD re-experiencing symptoms. Unlike Paulik, Steel, and Arntz (2019), who found further reductions in weekly trauma-related intrusions from post-treatment to 3-month follow up, we found medium and large increases in weekly trauma-related intrusions and the severity of PTSD re-experiencing symptoms, respectively, across these timepoints. Although this suggests that gains in PTSD re-experiencing symptoms were not fully maintained at 3-month follow-up, the overall reduction in re-experiencing symptoms at follow-up was nevertheless very large.

The analysis of the weekly measures may support a causal effect of treatment by showing no significant change during baseline, or even deterioration, and significant improvements during the treatment phase, with improvements still being significant at 3-month follow-up in all but one of the weekly measures. The exception was the frequency of positive voices that did not show lasting changes in frequency. Although scant attention has been given to the change in voice content (Salt et al. 2024), this finding is consistent with a recent clinical open trial of CBT in transdiagnostic voice hearers, which

found only changes in negative but not positive voice content (Brand, Badcock, and Paulik 2022). In general, weekly measures tended to show larger effects at posttreatment than at 3-month follow-up. A possible reason might be that for many participants only a limited number of trauma memories could be addressed, which deviates from other treatments that are typically longer, and/or the deviation from 'usual' ImRs where later in treatment patients do the rescripting themselves, which has been found to foster autonomy and empowerment and helps participants to continue using the technique after treatment (Twardawski et al. 2021). A recommendation for future clinical studies that are not so restricted in the number of sessions being delivered would be to offer these phase two and three of Arntz and Weertman's (1999) protocol and/or investigate the potential benefits of delivering ImRs in the broader context of schema mode therapy (Bartulovich et al. 2024).

#### 4.1 | Limitations and Future Research Directions

There are several limitations to this study that need to be addressed in future research. The psychometrics of the single-item weekly measures are unknown, which restricts confident interpretation of weekly data. Future studies may provide stronger evidence of weekly change trajectories by using valid and reliable full-scale measures but should consider the risk of participant burden and limit the volume of assessments were possible. The predominance of women in the sample may limit the generalisability of the findings to men. Nevertheless, the sample is considerably representative, given that compared to men, women are more likely to have PTSD symptoms and exposure to sexual trauma (Olf 2017), which is most strongly associated with trauma-related voices. Although this study design was consistent with current literature that emphasises a need to study voices separately from psychotic diagnoses and other psychotic-like symptoms (Carpenter and Regier 2016), the lack of diagnostic assessment may affect generalisability of these transdiagnostic findings to populations that

meet full diagnostic criteria for a psychotic spectrum disorder. Similarly, the generalisability of these findings to individuals who meet full diagnostic criteria for PTSD is uncertain. As trauma-related voice hearers often have complex PTSD, which is currently recognised in the International Classification of Diseases-11 (ICD-11; World Health Organization 2022), clinicians and future researchers may benefit from assessing for cPTSD in voice hearing clients in addition to traditional PTSD assessment tools.

Understanding the mechanisms maintaining trauma-related voice hearing and how ImRs influences such mechanisms is important for developing and refining interventions (Strachan, Paulik, and McEvoy 2022; Strachan et al. 2023). It may also be important to examine which aspects of ImRs are most important in producing changes in symptoms and potential mechanisms and how ImRs and CBT may differ in their modification of such mechanisms. Past ImRs studies (Ison et al. 2014; Paulik et al. 2019) and the current study used Arntz and Weertman's (1999) ImRs protocol, which focusses on modifying unmet needs. However, Clarke, Kelly, and Hardy (2022) used an ImRs protocol that focussed on modifying present-focused self-referential appraisals. Although the findings across studies are largely consistent, future studies could examine whether variations in the inclusions of specific content are associated with differential effects on symptoms and potential mechanisms.

Though non-significant change across multiple baseline assessments suggests that the observed effects are due to ImRs, the lack of a control group means that we cannot firmly conclude that ImRs produced these changes. Furthermore, although past studies suggest that CBT has limited effectiveness in treating voice hearing and PTSD in trauma-affected voice hearers (Brand et al. 2018; Paulik et al. 2019; Thomas et al. 2014), the lack of a direct comparison between CBT and ImRs with blinded outcome ratings precludes strong conclusions about the relative effectiveness of each intervention. Thus, randomised controlled trials are needed to compare the effectiveness of ImRs to current treatments.

## 5 | Conclusion

This study provides further evidence that ImRs is a safe and probably effective treatment for reducing trauma-related voice hearing and PTSD symptoms in people with a range of diagnoses; however, randomised controlled trials are needed to directly compare the efficacy of ImRs to existing CBT protocols.

### Acknowledgements

Open access publishing facilitated by Murdoch University, as part of the Wiley - Murdoch University agreement via the Council of Australian University Librarians.

### Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

### Endnote

<sup>1</sup>As voice hearers often report clinically significant symptoms of PTSD that do not meet full criteria for PTSD (Hardy, van de Giessen, and van den Berg 2019), trauma history included (a) exposure to DSM-5 Criterion A traumas (measured using the trauma screen from the Posttraumatic Diagnostic Scale for DSM-5; Foa et al. 2016) or (b) another event that caused significant distress (e.g., bullying). PTSD symptom severity or diagnosis of PTSD were not part of inclusion criteria; post-traumatic stress symptoms included any clinically significant symptoms associated with DSM-5 Criteria B to H (American Psychiatric Association 2013).

### References

- American Psychiatric Association. 2013. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. American Psychiatric Association.
- Arntz, A. 2011. "Imagery Rescripting for Personality Disorders." *Cognitive and Behavioral Practice* 18, no. 4: 466–481. <https://doi.org/10.1016/j.cbpra.2011.04.006>.
- Arntz, A. 2012. "Imagery Rescripting as a Therapeutic Technique: Review of Clinical Trials, Basic Studies, and Research Agenda." *Journal of Experimental Psychopathology* 3, no. 2: jep.024211. <https://doi.org/10.5127/jep.024211>.
- Arntz, A., and A. Weertman. 1999. "Treatment of Childhood Memories: Theory and Practice." *Behaviour Research and Therapy* 37, no. 8: 715–740. [https://doi.org/10.1016/S0005-7967\(98\)00173-9](https://doi.org/10.1016/S0005-7967(98)00173-9).
- Bailey, T., M. Alvarez-Jimenez, A. M. Garcia-Sanchez, C. Hulbert, E. Barlow, and S. Bendall. 2018. "Childhood Trauma Is Associated With Severity of Hallucinations and Delusions in Psychotic Disorders: A Systematic Review and Meta-Analysis." *Schizophrenia Bulletin* 44, no. 5: 1111–1122. <https://doi.org/10.1093/schbul/sbx161>.
- Bartulovich, A., P. Skeffington, P. Hastie, A. Arntz, C. Lee, and G. Paulik. 2024. "You're the Voice! Or Not? Try and Understand It: Exploring a Possible Relationship Between Voice Hearing Experiences and Schema Modes." *Psychology and Psychotherapy: Theory, Research and Practice*.
- Brand, R. M., J. C. Badcock, and G. Paulik. 2022. "Changes in Positive and Negative Voice Content in Cognitive-Behavioural Therapy for Distressing Voices." *Psychology and Psychotherapy* 95, no. 3: 807–819. <https://doi.org/10.1111/papt.12399>.
- Brand, R. M., C. McEnery, S. Rossell, S. Bendall, and N. Thomas. 2018. "Do Trauma-Focussed Psychological Interventions Have an Effect on Psychotic Symptoms? A Systematic Review and Meta-Analysis." *Schizophrenia Research* 195: 13–22. <https://doi.org/10.1016/j.schres.2017.08.037>.
- Carpenter, W. T., and D. Regier. 2016. "Diagnostic Categories: Provisional, Not Otherwise Classified, or Place-holder?" *Schizophrenia Bulletin* 42, no. 6: 1305–1306. <https://doi.org/10.1093/schbul/sbw127>.
- Chaudhury, S. 2010. "Hallucinations: Clinical Aspects and Management." *Industrial Psychiatry Journal* 19, no. 1: 5–12. <https://doi.org/10.4103/0972-6748.77625>.
- Clarke, R., R. Kelly, and A. Hardy. 2022. "A Randomised Multiple Baseline Case Series of a Novel Imagery Rescripting Protocol for Intrusive Trauma Memories in People With Psychosis." *Journal of Behavior Therapy and Experimental Psychiatry* 75: 101699. <https://doi.org/10.1016/j.jbtep.2021.101699>.
- de Bont, P. A. J. M., D. P. G. van Den Berg, B. M. van Der Vleugel, et al. 2015. "Predictive Validity of the Trauma Screening Questionnaire in Detecting Post-Traumatic Stress Disorder in Patients With Psychotic Disorders." *British Journal of Psychiatry: the Journal of Mental Science* 206, no. 5: 408–416. <https://doi.org/10.1192/bjp.bp.114.148486>.
- de Leede-Smith, S., and E. Barkus. 2013. "A Comprehensive Review of Auditory Verbal Hallucinations: Lifetime Prevalence, Correlates and

- Mechanisms in Healthy and Clinical Individuals." *Frontiers in Human Neuroscience* 7: 367. <https://doi.org/10.3389/fnhum.2013.00367>.
- Faul, F., E. Erdfelder, A.-G. Lang, and A. Buchner. 2007. "G\*Power 3: A Flexible Statistical Power Analysis Program for the Social, Behavioral, and Biomedical Sciences." *Behavior Research Methods* 39: 175–191. <https://doi.org/10.3758/BF03193146>.
- Foa, E. B., C. P. McLean, Y. Zang, et al. 2016. "Psychometric Properties of the Posttraumatic Diagnostic Scale for DSM–5 (PDS–5)." *Psychological Assessment* 28: 1166–1171. <https://doi.org/10.1037/pas0000258>.
- Hardy, A., R. Emsley, D. Freeman, et al. 2016. "Psychological Mechanisms Mediating Effects Between Trauma and Psychotic Symptoms: The Role of Affect Regulation, Intrusive Trauma Memory, Beliefs, and Depression." *Schizophrenia Bulletin* 42, no. suppl1: S34–S43. <https://doi.org/10.1093/schbul/sbv175>.
- Hardy, A., D. Fowler, D. Freeman, et al. 2005. "Trauma and Hallucinatory Experience in Psychosis." *The Journal of Nervous and Mental Disease* 193, no. 8: 501–507. <https://doi.org/10.1097/01.nmd.0000172480.56308.21>.
- Hardy, A., I. van de Giessen, and D. van den Berg. 2019. "Trauma, Posttraumatic Stress, and Psychosis." In *A Clinical Introduction to Psychosis: Foundations for Clinical Psychologists and Neuropsychologists*, edited by J. C. Badcock and G. Paulik, 223–243. Elsevier.
- Hawkins, N. G. B. A., R. W. P. Sanson-Fisher, A. P. Shakeshaft, C. P. D'Este, and L. W. P. Green. 2007. "The Multiple Baseline Design for Evaluating Population-Based Research." *American Journal of Preventive Medicine* 33, no. 2: 162–168. <https://doi.org/10.1016/j.amepre.2007.03.020>.
- Ison, R., L. Medoro, N. Keen, and E. Kuipers. 2014. "The use of Rescripting Imagery for People With Psychosis Who Hear Voices." *Behavioural and Cognitive Psychotherapy* 42, no. 2: 129–142. <https://doi.org/10.1017/S135246581300057X>.
- Jacobson, N. S., and P. Truax. 1992. "Clinical Significance: A Statistical Approach to Defining Meaningful Change in Psychotherapy Research." In *Methodological Issues & Strategies in Clinical Research*, edited by A. E. Kazdin, 631–648. American Psychological Association. <https://doi.org/10.1037/10109-042>.
- Kim, S. H., H. Y. Jung, S. S. Hwang, et al. 2010. "The Usefulness of a Self-Report Questionnaire Measuring Auditory Verbal Hallucinations." *Progress in Neuro-Psychopharmacology and Biological Psychiatry* 34, no. 6: 968–973. <https://doi.org/10.1016/j.pnpb.2010.05.005>.
- Larøi, F., I. E. Sommer, J. D. Blom, et al. 2012. "The Characteristic Features of Auditory Verbal Hallucinations in Clinical and Nonclinical Groups: State-Of-The-Art Overview and Future Directions." *Schizophrenia Bulletin* 38, no. 4: 724–733. <https://doi.org/10.1093/schbul/sbs061>.
- Llewellyn-Bennett, R., L. Bowman, and R. Bulbulia. 2016. "Post-Trial Follow-Up Methodology in Large Randomized Controlled Trials: A Systematic Review Protocol." *Systematic Reviews* 5, no. 1: 214. <https://doi.org/10.1186/s13643-016-0393-3>.
- Lovibond, S. H., and P. F. Lovibond. 1995. *Manual for the Depression-Anxiety Stress Scales*. 2nd ed. Psychology Foundation of Australia.
- Olf, M. 2017. "Sex and Gender Differences in Post-Traumatic Stress Disorder: An Update." *European Journal of Psychotraumatology* 8, no. suppl. 4: 1351204. <https://doi.org/10.1080/20008198.2017.1351204>.
- Paulik, G., M. Hayward, A. M. Jones, and J. C. Badcock. 2019. "Evaluating the "C" and "B" in Brief Cognitive Behaviour Therapy for Distressing Voices in Routine Clinical Practice in an Uncontrolled Study." *Clinical Psychology & Psychotherapy* 26, no. 6: 734–742. <https://doi.org/10.1002/cpp.2395>.
- Paulik, G., C. Steel, and A. Arntz. 2019. "Imagery Rescripting for the Treatment of Trauma in Voice Hearers: A Case Series." *Behavioural and Cognitive Psychotherapy* 47, no. 6: 709–725. <https://doi.org/10.1017/S1352465819000237>.
- Peach, N., M. Alvarez-Jimenez, S. J. Cropper, et al. 2020. "Trauma and the Content of Hallucinations and Post-Traumatic Intrusions in First-Episode Psychosis." *Psychology and Psychotherapy* 10: 223–241. <https://doi.org/10.1111/papt.12273>.
- Salt, E., P. Skeffington, C. Reddyhough, and G. Paulik. 2024. "Assessment of Voice Content for Voice Hearers: Psychometric Evaluation of the Perth Voice Content Questionnaire." *Clinical Psychology & Psychotherapy* 31, no. 3: e2991. <https://doi.org/10.1002/cpp.2991>.
- Strachan, L. P., G. Paulik, and P. M. McEvoy. 2022. "A Narrative Review of Psychological Theories of Post-Traumatic Stress Disorder, Voice Hearing, and Other Psychotic Symptoms." *Clinical Psychology & Psychotherapy* 29, no. 6: 1791–1811. <https://doi.org/10.1002/cpp.2754>.
- Strachan, L. P., G. Paulik, L. Roberts, and P. M. McEvoy. 2023. "Voice Hearers' Explanations of Trauma-Related Voices and Processes of Change Throughout Imagery Rescripting: A Qualitative Exploration." *Psychology and Psychotherapy: Theory, Research and Practice* 96: 982–998. <https://doi.org/10.1111/papt.12491>.
- Thomas, N., M. Hayward, E. Peters, et al. 2014. "Psychological Therapies for Auditory Hallucinations (Voices): Current Status and Key Directions for Future Research." *Schizophrenia Bulletin* 40, no. SUPPL. 4: S202–S212. <https://doi.org/10.1093/schbul/sbu037>.
- Tolmeijer, E., A. Hardy, A. Jongeneel, A. B. P. Staring, M. van der Gaag, and D. van den Berg. 2021. "Voice-Hearers' Beliefs About the Causes of Their Voices." *Psychiatry Research* 302: 113997. <https://doi.org/10.1016/j.psychres.2021.113997>.
- Twardawski, M., M. Gollwitzer, M. S. Altmüller, A. E. Kunze, and C. E. Wittekind. 2021. "Imagery Rescripting Helps Victims Cope With Experienced Injustice." *Zeitschrift für Psychologie* 229, no. 3: 178–184. <https://doi.org/10.1027/2151-2604/a000450>.
- van den Berg, D. P. G., P. A. J. M. de Bont, B. M. van der Vleugel, et al. 2015. "Prolonged Exposure vs Eye Movement Desensitization and Reprocessing vs Waiting List for Posttraumatic Stress Disorder in Patients With Psychotic Disorder: A Randomised Clinical Trial." *JAMA Psychiatry* 72: 259–267. <https://doi.org/10.1001/jamapsychiatry.2014.2637>.
- Van Lieshout, R. J., and J. O. Goldberg. 2007. "Quantifying Self-Reports of Auditory Verbal Hallucinations in Persons With Psychosis." *Canadian Journal of Behavioural Science* 39, no. 1: 73–77. <https://doi.org/10.1037/cjbs2007006>.
- Ventura, J., M. F. Green, A. Shaner, and R. P. Liberman. 1993. *Training and Quality Assurance With the Brief Psychiatric Rating Scale: "The Drift Busters"*. Vol. 3, 221–244. US: John Wiley & Sons.
- World Health Organization. 2022. "ICD-11: International Classification of Diseases (11<sup>th</sup> Revision)." <https://icd.who.int/>.
- Zalta, A. K. 2015. "Psychological Mechanisms of Effective Cognitive–Behavioral Treatments for PTSD." *Current Psychiatry Reports* 17: 23. <https://doi.org/10.1007/s11920-015-0560-6>.

### Supporting Information

Additional supporting information can be found online in the Supporting Information section.