

# Technology-enhanced psychological assessment and treatment of distressing auditory hallucinations: a systematic review

Emma O'Neill, Molly Bird and Simon Riches

*King's College London, Institute of Psychiatry, Psychology and Neuroscience, London, UK and South London and Maudsley NHS Foundation Trust, London, UK*

## Abstract

**Purpose** – Auditory hallucinations (“hearing voices”) are a relatively common experience, which is often highly distressing and debilitating. As mental health services are under increasing pressures, services have witnessed a transformative shift with the integration of technology into psychological care. This study aims to narratively synthesise evidence of technology-enhanced psychological assessment and treatment of distressing voices (PROSPERO 393831).

**Design/methodology/approach** – This review was carried out according to the preferred reporting items for systematic reviews and meta-analyses. Embase, MEDLINE, PsycINFO and Web of Science were searched until 30th May 2023. The Effective Public Health Practice Project (EPHPP) tool assessed methodological quality of studies.

**Findings** – Searching identified 9,254 titles. Fourteen studies (two assessment studies, twelve treatment studies, published 2010–2022,  $n = 1,578$ ) were included in the review. Most studies were conducted in the UK, the USA or Canada. Technologies included avatar therapy, mobile apps, virtual reality, a computerised Web-based programme and a mobile-assisted treatment. Overall, technology-enhanced psychological assessments and treatments appear feasible, acceptable and effective, with avatar therapy the most used intervention. EPHPP ratings were “strong” ( $n = 8$ ), “moderate” ( $n = 5$ ) and “weak” ( $n = 1$ ).

**Originality/value** – To the best of the authors’ knowledge, this is the first systematic review to investigate these technologies, specifically for distressing voices. Despite the relatively small number of studies, findings offer promising evidence for the clinical benefits of these technologies for enhancing mental health care for individuals with distressing voices. More high-quality research on a wider range of technologies is warranted.

**Keywords** Avatar therapy, Auditory hallucinations, Cognitive behavioural therapy, Digital mental health, Psychosis, Virtual reality

**Paper type** Research paper

## Introduction

The experience of hearing voices is a relatively common experience in the general population, with an estimated one in ten adults reporting the experience across their lifetime (Majjer *et al.*, 2018). Voices are a common feature of psychotic disorders, although evidence suggests that it is a transdiagnostic experience, presenting in a range of other clinical groups, including post-traumatic stress disorder (PTSD), eating disorders and emotionally unstable personality disorder (Schutte *et al.*, 2020).



It can be a highly distressing and debilitating experience which can have a significant impact on the individual, such as high levels of distress and functional disability. Consistent with the continuum model of psychosis, the experience of voice hearing extends beyond clinical groups to non-clinical groups in the general population, where there is not necessarily associated distress or a “need for care” (Baumeister *et al.*, 2017).

The UK National Health Service (NHS) is under increasingly significant pressure, as demand for services vastly exceeds resources and there are often significant waiting times to access psychological treatment, such as cognitive behavioural therapy for psychosis (CBTp). Further, meta-analyses have consistently reported that CBTp offers moderate effects on psychotic symptoms at best (Wykes *et al.*, 2008; Jauhar *et al.*, 2014; Turner *et al.*, 2014). There is, therefore, a need to develop more accessible psychological treatments for individuals who experience distressing voices.

There have been significant developments in the therapeutic use of computers, mobile phones and virtual reality (VR) technologies for the treatment of mental health conditions, with increased opportunities for integrating innovative technologies into clinical practice (Torous *et al.*, 2021). Digital health technologies may greatly enhance delivery of evidence-based psychological treatments and reduce staff burden in stretched mental health services (Steinhubl *et al.*, 2013). Another potential merit is that they can be highly accessible, allowing clients to access mental health support remotely and in-between face-to-face therapy sessions or to act as an anonymous, destigmatising steppingstone to traditional in-person treatment (Bond *et al.*, 2023). In the context of these innovations in digital mental health, it is thought that some technologies merely “enable” the delivery of psychological assessments and treatments, such as therapy sessions delivered via video call; but other technologies can have an “enhancing” function, whereby the technology offers an aspect to the intervention that could not be achieved without the technology (Bond *et al.*, 2023).

The evolving evidence on the use of technology-enhanced psychological assessments and treatments has predominantly focused on common mental health conditions, such as anxiety and depression (Baños *et al.*, 2022). However, evidence from systematic reviews indicates that these technologies can effectively improve clinical outcomes and social functioning for individuals with psychosis (Bell *et al.*, 2017; Bonet *et al.*, 2017; Clarke *et al.*, 2019; Riches *et al.*, 2021). A recent systematic review highlighted 21 digital health technologies for people with psychosis, which incorporated a mixture of computer-assisted, avatar and phone app-based approaches (Clarke *et al.*, 2019). Findings provided preliminary evidence for their effectiveness in reducing psychotic symptoms, with avatar-based therapies appearing to hold the most promise. Studies have consistently indicated that individuals with psychosis are interested in using digital health interventions and can use them without adverse effects (Craig *et al.*, 2018; Bucci *et al.*, 2018; Maroño Souto *et al.*, 2018). In one study, interviews with early psychosis service users revealed that they were largely positive about the potential use of digital health interventions in supporting and managing their mental health difficulties (Bucci *et al.*, 2018). Among those interviewed, there was agreement that mobile technology is an acceptable tool to access support for mental health problems.

Despite this emerging evidence for digital health interventions for people with psychosis, previous reviews have typically targeted a range of psychotic symptoms or psychosis more broadly, with limited studies targeting specific symptoms, such as hearing voices. Therefore, less is known about what technology-enhanced assessment and treatment is available for people who hear voices. As distressing voices are a key symptom of psychotic disorders but are also associated with high levels of distress and impairment in other clinical presentations and non-clinical groups can also report distress (Toh *et al.*, 2022; Connell *et al.*,

2019), the current review investigated technology-enhanced psychological assessment and treatment of distressing voices for both clinical and non-clinical groups. It investigated what technologies are available, and evaluated their feasibility, acceptability and effectiveness for use in psychological assessment and treatment.

### Methods

This review was carried out according to the preferred reporting items for systemic reviews and meta-analyses (PRISMA) (Moher *et al.*, 2009). This review was pre-registered on PROSPERO (CRD42023393831).

Database searches were completed on 30th May 2023 using databases Embase, MEDLINE, PsycINFO and Web of Science. Searches were completed separately for each database, using truncations and using the abstract, keyword and title search fields. The following search terms were used: voice\* OR auditory hallucinations OR AVH OR exp auditory hallucinations/AND technol\* OR internet OR Web\* OR computer\* OR online OR digital OR app OR smartphone or virtual real\* OR VR OR virtual character\* OR VCs OR virtual environ\* OR augmented reality OR avatar\* or ehealth OR e-health OR mhealth OR m-health or wearable\* or artificial intelligence OR AI OR exp digital technology/AND psychotherap\* or psycholog\* OR therap\* OR psychological assess\* OR psychological treat\* OR intervent\* OR self-help OR exp/clinical psychology. The “explode” function was used for the following subject headings: “digital technology”, “voices” and “clinical psychology” across the OVID databases (PsycINFO, MEDLINE and Embase), to search for more specific terms within the broader headings.

Two raters (EO and MB) independently ran all searches on each database, to check for consistency in results. Reference lists of previous systematic reviews in the relevant area were reviewed. Database searches were limited by Human and English language publications.

### *Inclusion and exclusion criteria*

Studies were included in the review if they were empirical, published in a peer-review journal, written in English, used quantitative research methods of any design, had a sample size of at least five, tested any populations who experienced distressing voices, used a voices-specific outcome measure and tested a technology-enhanced psychological assessment and/or treatment. Specifically, technologies needed to enhance (rather than merely enable) the psychological assessment or treatment. Studies were excluded if they were abstracts, conference proceedings, dissertations, non-empirical, reviews or used only qualitative methods. Studies were also excluded if they investigated only participants’ interest or willingness to receive a technology-enhanced treatment.

### *Data extraction*

Studies were extracted and downloaded onto reference management software Endnote and an Excel spreadsheet for screening purposes. The above inclusion and exclusion criteria were used to review all abstracts and titles by one researcher (EO), with 20% of these search results independently screened by another researcher (MB). The same procedure was completed for the screening of full-text papers, with one independent rater (EO) reviewing all included full texts against the inclusion and exclusion criteria and 20% of the full text papers were then screened by another researcher (MB). In the event of any discrepancies or disagreement between researchers, studies were discussed between the research team (EO, MB, SR) until discrepancies were resolved.

For each included publication, data on the following information was extracted: study title, year of study, country of study, study design, sample size, population sampled (clinical

---

or non-clinical and demographics, such as age, gender and ethnicity), voice-related outcome measure, technology studied, procedure for using the technology, and key findings.

### *Quality assessment*

All quality ratings were carried out by two independent reviewers (EO and MB), under the supervision of a senior clinical researcher (SR), using the Effective Public Health Practice Project (EPHPP) tool (Ciliska *et al.*, 1998). The EPHPP has good content and construct validity, and inter-rater reliability (Thomas *et al.*, 2004) and can provide consistent quality ratings for a range of study designs. EPHPP's six subscales (selection bias, study design, confounders, blinding, data collection methods and withdrawals and drop-outs) are given a rating of "strong", "moderate" or "weak". A global rating for each study is then calculated. Studies receive a global rating of "strong" if there are no weak subscale ratings, "moderate" if there is one weak subscale rating, and "weak" if there are two or more weak subscale ratings. EPHPP reclassifies randomised controlled trials (RCTs) as controlled clinical trials (CCTs) if studies do not report information on the method of randomisation. For the confounders subscale, studies were rated as "N/A" if the study sampled one group. Discrepancies in ratings were discussed between researchers (EO, MB, SR) and studies were re-evaluated until consensus was reached.

### *Narrative synthesis*

The review used a narrative approach to synthesise findings, using the synthesis without meta-analysis reporting guidelines for systematic reviews (Campbell *et al.*, 2020). Studies were organised into assessment and treatment studies, with findings then organised in terms of feasibility, acceptability and effectiveness. Feasibility was evaluated based on the reports of accessibility, whether the researchers were able to deliver the assessment or treatment, retention rates and rates of withdrawals and dropouts. Acceptability was evaluated using any reports of adverse effects, user experience (including any acceptability-related measures) and qualitative feedback from participants. To assess the effectiveness of the technologies, study outcomes were reviewed, specifically the impact on distressing voices.

## **Results**

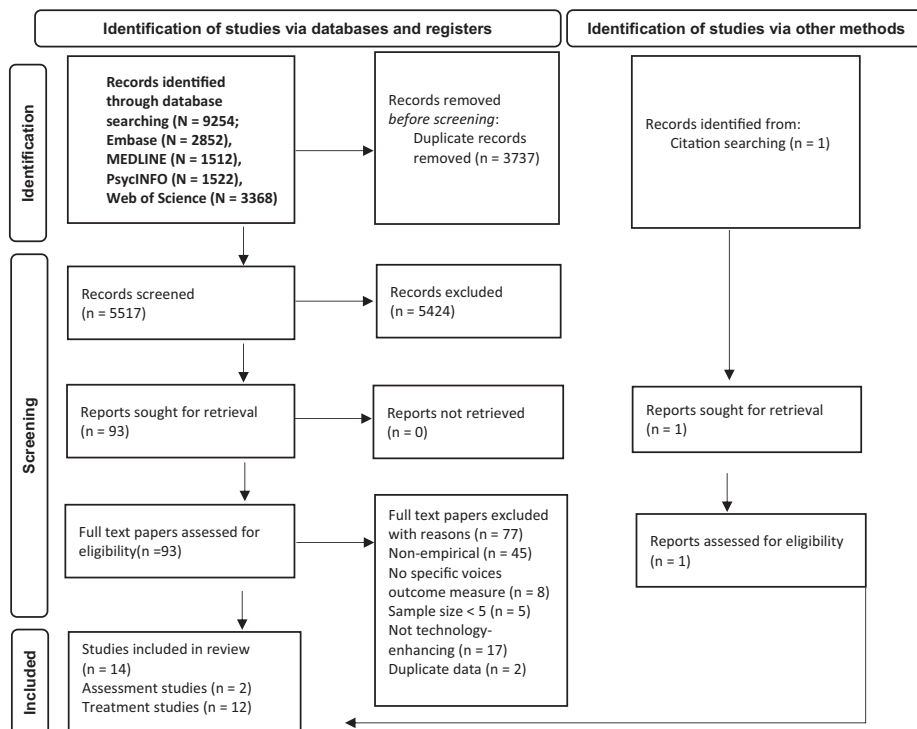
Database searching on 30th May 2023 identified a total of 9,254 titles. There were 5,517 titles identified for screening after de-duplicating and removing books and book sections in Endnote, of which 5,424 were excluded following title/abstract screening, leaving 93 full-text articles for assessment. Full screening process and reasons for study exclusions are detailed in [Figure 1](#).

### *Study characteristics*

A total of 14 studies, published between 2010 and 2022, were included in the review. Studies were conducted in the UK ( $n = 5$ ), the USA ( $n = 4$ ), Canada ( $n = 3$ ), the Netherlands ( $n = 1$ ) and Australia ( $n = 1$ ). Full details of study characteristics are shown in [Table 1](#).

Studies comprised two assessment studies and 12 treatment studies. Out of the two assessment studies, one study was a CCT (Stinson *et al.*, 2010) and the other was a cohort study (Cardi *et al.*, 2022). Out of the 12 treatment studies, there were six cohort studies, three CCTs and three RCTs.

The two assessment studies assessed voices using computerised avatar therapy (Cardi *et al.*, 2022) and VR (Stinson *et al.*, 2010). The treatment studies investigated the use of avatar therapy



**Figure 1.**  
PRISMA 2020 flow  
diagram

**Source:** Created by authors

[either with computerised software ( $n = 4$ ) or in VR ( $n = 3$ ), mobile apps ( $n = 3$ ), a computerised Web-based programme ( $n = 2$ ) and a mobile device-assisted treatment ( $n = 1$ ).

The included studies had a total of 1,578 participants, with sample sizes ranging from 10 to 1,048. All studies sampled participants aged 14 years and older, with an overall mean age of 40.8 years. All the studies reported the gender of participants, with most studies having a higher proportion of male participants than female participants, although one of the UK studies on the eating disorder voice sampled only female participants (Cardi *et al.*, 2022). Of the nine studies that reported ethnicity data, most tested a majority of participants from a White ethnic background.

All studies, except from one (Jongeneel *et al.*, 2022), sampled clinical groups, mostly including those with a schizophrenia spectrum or psychosis condition. One study (Cardi *et al.*, 2022) sampled 39 participants with a diagnosis of anorexia nervosa or in remission. Other clinical participants included those with a diagnosis of PTSD, major depressive disorder and bipolar disorder. The only study that included non-clinical participants was a Netherlands app study, which included any app user who self-reported hearing voices (Jongeneel *et al.*, 2022).

#### *Assessment studies*

There were two assessment studies:

*Avatar computerised software* ( $n = 1$ ). One UK assessment study (Cardi *et al.*, 2022) sampled 39 females with a diagnosis of anorexia nervosa or in remission to assess their eating disorder voice. Participants completed a baseline eating disorder voice assessment.

Study	Country	Study design	Sample	N, gender	Mean age (SD), ethnicity	Technology	Participants' task	Voices measures	Findings
Assessment studies (n = 2) Stinson <i>et al.</i> (2010)	The UK	CCT	Adults with auditory hallucinations occurring at least once a day, which occurred in social contexts recruited from South London and Maudsley NHS Foundation Trust or internet website for voice-hearers	30 (10F)	Mean age = 42.4 (9.7) no ethnicity data	VR	Enter four-min VR tube journey while being instructed to focus on cognitive antecedent for their voice. Afterwards, describe any voices experienced while in VR and then complete the TVRS related to the time in VR.	PSYRATS-AH, TVRS	No reported dropouts or withdrawals from the study. VR did not induce motion sickness. Anxiety and heart rate levels remained consistent from pre-VR to post-VR. 93% of participants reported not experiencing unwanted thoughts about the VR in the following week. Ten participants heard voices in VR, compared to eight in the control group.
Cardi <i>et al.</i> (2022)	The UK	Cohort study	Individuals with a current DSM-5 diagnosis of anorexia nervosa or in remission	39 (39F)	Mean age (anorexia nervosa) = 23.1 (6.9), mean age (in remission) = 25.3 (4.0) no ethnicity data	Avatar computerised software	Complete baseline assessment on the eating disorder voice, then create a digital representation (visual and auditory) of the eating disorder voice using a computerised software. Exposure to the avatar's face they created, which spoke the statement the participant reported as the most distressing, then complete the avatar's creation feedback	Eating disorder voice assessment (five closed questions) Avatar's creation feedback (four visual analogue scales)	No reported dropouts or withdrawals from the study. Almost all participants identified the eating disorder voice as human and internal. Almost 90% of individuals reported some level of distress in response to exposure to the avatar. None of the participants across groups reported feeling "very distressed" and all said they would be willing to be re-exposed to their eating disorder avatar in the future.

(continued)

**Table 1.** Characteristics of studies on technology-enhanced psychological assessments and treatments for distressing voices

Table 1.

Study	Country	Study design	Sample	N, gender	Mean age (SD), ethnicity	Technology	Participants' task	Voices measures	Findings
<i>Treatment studies (n = 12)</i> <b>Granholtm et al. (2012)</b>	The USA	Cohort study	Adults (aged over 18) living in the community with schizophrenia or schizoaffective disorder	42 (13F)	Mean age = 48.7 (9.1) No ethnicity data	<i>Mobile phone-assisted treatment</i> Mobile assessment and treatment for schizophrenia (MATS)	Initial assessments gathered information from participants to create personalised thought-challenging messages. Subsequently, four text messages were sent to participants daily, with each message targeting voices	Daily monitoring assessment question – "have you been bothered by voices?"	13 out of 55 were "non-completers" due to them not sending any valid messages or stopped sending valid messages within two weeks. For completers, the valid response rate for the assessment question was M = 86%, which was followed by the voices intervention. Majority (86%) of phones returned intact. Participants reporting finding the mobile intervention moderately to very helpful. Significant reduction in likelihood of being bothered by voices Data was available for 31 (91%) participants at post-treatment, with three participants lost to follow-up and no withdrawals. 13 out of 17 participants completed all four sessions. All participants agreed that they would recommend it to other people who hear voices. Verbatim open feedback included: It helped me to control my voices and to make me feel better about myself. Significant reductions in PSYRATS-AH total scores
<b>Bell et al. (2020)</b>	Australia	Single-blind, parallel group, pilot RCT	Adults (18+) from a specialist voices clinic, clinical services and consumer groups, who experienced current, frequent and distressing voices for at least six months	34 (19F)	Mean age (experimental group) = 39.12 (10.64), mean age (TAU) = 42.59 (10.64) Ethnicity (experimental group) = 70.6% Australian, 11.6% British or Irish, 5.9% New Zealander, 5.9% Greek and 5.9% Other; ethnicity (TAU) = 94.1% Australian, 5.9% Greek	<i>App</i> SAVvy, a brief in-person therapy blended with an app	In-person introduction and training session for app usage, six days of EMA monitoring of voices via an app, followed by three in-person therapy sessions on developing coping strategies, with EMI (coping reminders, surveys and feedback) via app in between sessions	PSYRATS-AH	

(continued)

Study	Country	Study design	Sample	N, gender	Mean age (SD), ethnicity	Technology	Participants' task	Voices measures	Findings
<a href="#">Buck et al. (2022)</a>	The USA	Cohort study	Individuals receiving treatment from an outpatient psychosocial rehabilitation clinic – with serious and chronic mental illness with current or past psychotic symptoms	17 (5F)	Mean age = 55.12 (13.02) Ethnicity = 11 (65%) White, 3 (18%) Black or African American, 2 (12%) Asian, 1 (6%) American Indian or Alaskan Native	<i>App</i> FOCUS mHealth intervention	Access brief, pre-programmed interventions on smartphone on demand, as well as receive prompts to remind users to complete assessments and receive interventions	Hamilton programme for voices questionnaire	On average, participants interacted with the app on 19.29 of 30 access days. Most participants (16, 94%) reported that they would recommend FOCUS to a friend. Participants reported that they liked that FOCUS was consistently available to them and that they were able to access helpful tools in the moment. Small positive effects were detected for participants in the severity of voices, but no significant reductions (analyses were underpowered to detect significant clinical effects)
<a href="#">Leff et al. (2013)</a>	The UK	Randomised, partial crossover trial, CCT	Clients of community mental health teams in a London mental health trust. Hearing persecutory voices for at least six months	26 (10F)	No age or ethnicity data	Avatar therapy Computerised software – Facegen Modeler version 3.5.1 for Windows, amresoft Real-time LipSync SDK 4.0.0.0 for Windows	Create an avatar of their persecutor and engage in dialogue with the avatar	PSYRATS-AH BAVQR omnipotent and malevolence subscales	Around a third of participants dropped out. One participant could not tolerate seeing the face of the avatar, which resembled the perpetrator from her sexual abuse. She was able to talk with the avatar when the face was deleted. Two participants heard multiple voices and could not concentrate on the avatar because the

(continued)

Table 1.



Table 1.

Study	Country	Study design	Sample	N, gender	Mean age (SD), ethnicity	Technology	Participants' task	Voices measures	Findings
<b>Gottlieb <i>et al.</i> (2013)</b>	The USA	Cohort pilot study	Adults with a psychotic disorder who were receiving outpatients mental health services at a clinic in the USA	17 (6F)	Mean age = 40.10 (13.63) Ethnicity = 57% Caucasian, 38% African American, 5% Asian	<i>Computerised, Web- based CBTp program</i> CwV	Supported by a research assistant, work through the CwV intervention, a highly interactive, computerised, Web- based CBTp programme	PSYRATS- AH, BAVQR	other voices spoke too loudly at the same time. Significant reductions in total PSRATS score (AH) and BAVQR combined score of omnipotence and malevolence of the voice In total, 17 out of 21 participants who completed baseline assessments were exposed to the CwV programme (i.e. completed at least six lessons). A total of 82% rated the programme overall as "Very good/ helpful". Significant reductions from baseline to posttreatment in several measures of auditory hallucinations, including overall severity and the perception of voices as an "outside entity" Majority (15 out of 19) participants randomised to CwV completed all 10 modules. All participants rated the programme as "very useful" (66%) or "useful" (63%) and approximately 80% stated that they would
<b>Gottlieb <i>et al.</i> (2013)</b>	The USA	RCT	Adults with a psychotic disorder who were receiving outpatients mental health services at a clinic in the USA	37 (14F)	Mean age (experimental group) = 43.79 (13.16), mean age (UC group) = 40.28 (11.69), ethnicity (experimental) = 15 (78.9%) White, 2 (10.5%) Black, 2 (10.5%) more than one race, ethnicity (UC) = 10 (65.6%) White, 3	<i>Computerised, Web- based CBTp program</i> "Coping with voices"	Supported by a research assistant, work through the CwV intervention, a highly interactive, computerised, Web- based CBTp programme	PSYRATS AH	(continued)

Study	Country	Study design	Sample	N, gender	Mean age (SD), ethnicity	Technology	Participants' task	Voices measures	Findings
Craig <i>et al.</i> (2018)	The UK	Single-blind RCT	Adults (aged 18–65) with a clinical diagnosis of a schizophrenia spectrum or affective disorders with enduring auditory verbal hallucinations, despite treatment	150 (48F)	(16.7%) Black, 2 (11.1%) Asian, 2 (11.1%) more than one race, 1 (5.6%) other Mean age = 42.7 (10.7) Ethnicity = 38 (39%) White British, 26 (17%) Black British, 15 (10%) Black Caribbean, 13 (9%) Black African, 4 (3%) Asian Indian, 1 (1%) Asian Chinese, 33 (22%) other	<i>Avatar therapy</i> Computerised avatar therapy software	Create a computerised representation (avatar) of their main voice, engage in dialogue with the avatar	PSYRATS-AH BAVQ-R; perceived malevolence, omnipotence and PSYRATS-AH total scores, as well as PSYRATS subscales of voice frequency and VAAAS; acceptance and action subscale VPDS; power and assertiveness subscales	recommend the programme to a friend or relative. Significant reduction in PSYRATS-AH and BPRS-AH Majority of participants (53 out of 75) completed therapy. No adverse effects attributed to the AVATAR therapy. Significant reductions in PSYRATS-AH total scores, as well as PSYRATS subscales of voice frequency and VAAAS; distress, BAVQ-R – omnipotence, VAAAS-acceptance and VAAAS-action scores. At 24 weeks of follow-up, improvements in scores on PSYRATS-AH, BAVQ and VAAAS in the AVATAR group were maintained
Di Sert <i>et al.</i> (2018)	Canada	Randomised, partial crossover trial, CCT	Patients with treatment-resistant schizophrenia or schizoaffective disorder	15 (6F)	Mean age = 42.9 (24.62) Ethnicity = 13 (86.7%) Caucasian, 2 (13.3%) other minority	<i>Avatar therapy</i> VR – Unity 3D game engine, Morph3D character system, voice transformer – Roland AIRA VT-3, SALSA with Random Eyes Unity 3D extension, VR-Samsung Gear VR head mounted display and Samsung Galaxy S6 smartphone	Create an avatar of their persecutor and immerse into VR setting to engage in dialogue with the avatar	PSYRATS-AH, BAVQ	Out of 19 participants, four dropped out of avatar therapy due to anxiety after the first therapeutic session and a lack of engagement in the therapy model. Participants rated their avatars credible enough to make them feel in presence of their persecutor. No participants re-hospitalised during the trial. Significant improvement in AVH severity

(continued)

Table 1.

Table 1.

Study	Country	Study design	Sample	N, gender	Mean age (SD), ethnicity	Technology	Participants' task	Voices measures	Findings
Rus-Calafell <i>et al.</i> (2020)	The UK	Cohort study	All participants allocated to AVATAR therapy as part of the AVATAR trial. Adults (over 18) with troubling auditory hallucinations and primary diagnosis of non-organic psychosis	39 (9F)	Mean age = 43.87 (9.33) Ethnicity = 13 (33%) White British, 5 (13%) Black British, 3 (8%) Black Caribbean, 5 (13%) Black African, 3 (8%) Asian Indian, 10 (25%) other	<i>Avatar therapy</i> Computerised avatar therapy software	Create a computerised representation (avatar) engage in dialogue with the avatar	PSYRAYS- AH	Frequency of voices reduction related to one predictor: the interaction between anxiety and sense of presence. Mid to high levels of sense of presence reported were consistently across therapy sessions, indicating that the avatar dialogue was effective in its aim of delivering valid simulation of the person's voice Participants reported that VRT helped to embody their voices and make their experience come to life by enabling a direct discussion with their voice (e.g. "the avatar was truly there" and they "had to face it"). Combination of CBT and VRT associated with significant reductions in voice severity, frequency and loudness Out of the 74 participants, nine withdrew, with reasons including lack of motivation, not wanting to reduce their voices and moving away. No participants re-
Dellazizzo <i>et al.</i> (2020)	Canada	Cohort study	Adult patients (aged over 18) with treatment resistant auditory verbal hallucinations and schizophrenia or schizoaffective disorder	10 (2F)	Mean age = 43.4 (14.6) Ethnicity = 10 (100%) Caucasian	<i>Avatar therapy</i> VR	Create a computerised representation (avatar) of their most distressing voice, immerse into a VR setting to engage in dialogue with the avatar	PSYRAYS- AH, BAVQR	Participants reported that VRT helped to embody their voices and make their experience come to life by enabling a direct discussion with their voice (e.g. "the avatar was truly there" and they "had to face it"). Combination of CBT and VRT associated with significant reductions in voice severity, frequency and loudness Out of the 74 participants, nine withdrew, with reasons including lack of motivation, not wanting to reduce their voices and moving away. No participants re-
Dellazizzo <i>et al.</i> (2021)	Canada	Pilot randomised comparative trial, CCT	Patients with treatment-resistant schizophrenia	74 (18F)	Mean age (VR) = 43.6 (12.0), mean age (CBT) = 41.4 (13.4) Ethnicity = 82.2% Caucasian, 17.8% "visible minorities"	<i>Avatar therapy</i> VR	Create a computerised representation (avatar) of their most distressing voice, immerse into VR setting to engage in dialogue with the avatar	PSYRAYS- AH, BAVQR	Participants reported that VRT helped to embody their voices and make their experience come to life by enabling a direct discussion with their voice (e.g. "the avatar was truly there" and they "had to face it"). Combination of CBT and VRT associated with significant reductions in voice severity, frequency and loudness Out of the 74 participants, nine withdrew, with reasons including lack of motivation, not wanting to reduce their voices and moving away. No participants re-

(continued)

Study	Country	Study design	Sample	N, gender	Mean age (SD), ethnicity	Technology	Participants' task	Voices measures	Findings
Jongeneel <i>et al.</i> (2022)	The Netherlands	Naturalistic cohort study	Any person in the Netherlands who downloaded Tenstem app (which is disseminated in Dutch networks of mental health professionals and individuals, and it is advertised on multiple websites) and reported hearing voices	1,048 (789F)	Mean age = 35.34 (14.03) no ethnicity data	App "Tenstem"	Select the silencing or challenging mode and then play relevant games for as long as they like	Voice distress or emotionality and vividness of voice memories rated on a 1-7 Likert scale	hospitalised during the trial. Most participants found their intervention to be adequate. A total of 37.5% reported that the intervention was stressful at first, but once they had overcome the initial exposure to anxiety, they enjoyed their experience and found it to be interesting. Significant improvements in the severity of voice and in beliefs about voices. Out of 3,609 users, 2,561 (71% users were excluded due to not reaching level 2 or playing <15 times. Significant reductions in voice-hearing distress and emotionality and vividness of voice memories

**Notes:** Demographics: F = female; DSM = the Diagnostic and Statistical Manual of Mental Disorders; TAU = treatment as usual; UC = usual care; CBT = cognitive behavioural therapy; VR = virtual reality. Measures and findings = PSYRATS-AH = the Psychotic Symptom Rating Scale – Auditory Hallucinations; TVPS = Topography of Voices Rating Scale; BAVQR = Beliefs about Voices Questionnaire – Revised; VAAS = Voices Acceptance and Action Scale; VPDS = Voices Power Differential Scale; AVH = auditory verbal hallucinations; EMI = ecological momentary intervention; CwV = coping with voices

**Source:** Created by authors

Table 1.

They were then guided by a research assistant to use a computerised software to create a digital representation of the eating disorder voice (i.e. an avatar's face) and were then exposed to this avatar, which spoke the most distressing statement their eating disorder voice says to them. Finally, they assessed closeness of match between their eating disorder voice and the avatar.

*Virtual reality (n = 1).* One assessment study used a virtual London underground to explore the occurrence of voices during VR (Stinson *et al.*, 2010). Fifteen participants identified cognitive antecedents to their voices and then experienced a 4-min VR tube journey with computer-generated commuters of both sexes and several ethnicities. The environment was designed to be neutral and non-threatening. During VR, participants were prompted to focus on the cognitive antecedent to their voice and after being in VR, participants' voices were assessed in relation to their time in VR.

### *Treatment studies*

There were 12 treatment studies:

*Avatar therapy (N = 6).* There were three studies investigating avatar therapy on a two-dimensional (2D) screen using computerised software and three studies using VR.

Two of the UK trials sampled participants with enduring auditory hallucinations to investigate avatar therapy using computerised software to develop the avatar (Leff *et al.*, 2013; Craig *et al.*, 2018). The therapist sat in an adjacent room and could view the 2D computer screen where the avatar was presented, while participants were encouraged to engage in dialogue with the avatar and stand up to them. The main aim was to enable participants to challenge their beliefs about the power of the voices and gain more control over them. A cohort study (Rus-Calafell *et al.*, 2020) formed part of the larger avatar trial (Craig *et al.*, 2018) and sampled only those in the AVATAR arm of the trial, to investigate the impact of voice presence on AVATAR therapy outcomes.

Three studies in Canada investigated avatar therapy using *immersive* VR (Du Sert *et al.*, 2018; Dellazizzo *et al.*, 2020, 2021), which involved using a head-mounted display to deliver the therapy. The treatment process was similar to the previous avatar therapy studies, but instead of a 2D computerised avatar, participants were immersed in a virtual environment and engaged with the avatar in three-dimensional (3D). The avatar was standing in the dark, seen from a first-person perspective.

*Mobile apps (N = 3).* Three studies investigated mobile apps (Bell *et al.*, 2020; Buck *et al.*, 2022; Jongeneel *et al.*, 2022). The Netherlands study investigated Temstem, an app made for and with voice-hearers, which is freely available in the Netherlands (Jongeneel *et al.*, 2022). When logged on to the app, an avatar provides psychoeducation about voices and the user selects a game to play, which they can play as many times as they wish. Users rated their voice distress or emotionality and vividness of a voice memory before and after playing the game, with the aim being to reduce these experiences.

Another study investigated FOCUS mHealth treatment for Veterans with Serious Mental Illness in the USA (Buck *et al.*, 2022). The mHealth treatment consists of three components: a mobile app, a clinician dashboard and a mHealth support specialist. The mobile app includes brief, pre-programmed self-management treatments accessed on demand on a smartphone, as well as via prompts. Users could access tailored treatment following completion of a brief ecological assessments (EMA) or more generic treatment via a toolbox.

An Australian study investigated "SAVVY", a brief face-to-face therapy blended with an ecological momentary assessment and treatment (EMA/EMI) via a smartphone app (Bell *et al.*, 2020). The treatment involved four face-to-face therapy sessions, with EMA

(completing a survey) or EMI (receiving personalised coping reminders) via a smartphone app in between those sessions.

*Computerised, Web-based cognitive behavioural therapy for psychosis programme (n = 2).* Two treatment studies investigated “Coping with Voices” (CwV) – a highly interactive, computerised, Web-based CBTp programme with a sample of adults with a psychotic disorder in the USA (Gottlieb *et al.*, 2013, 2017). The programme consisted of ten lessons and included animated tutorials and videos, exercises and interactive games, tracking of symptoms and goals and independent skills practice worksheets. The main goal of the programme is to reduce distress, preoccupation and day-to-day interference of voices, using CBT-based skills. Sessions took place on a laptop, with a research assistant situated in an adjacent room, who set up the participant on the laptop, reviewed progress, provided basic technical support, reviewed homework worksheets, answered questions, scheduled the next appointment and made an appointment reminder call.

*Mobile device-assisted treatments (n = 1).* One cohort study sampled adults with schizophrenia or schizoaffective disorder living in the community in the USA to investigate a mobile-device assisted treatment (Granholtm *et al.*, 2012). Participants completed an in-person interview to create personalised thought-challenging messages and then they received daily text messages targeting the voices. The text messages included two multi-choice assessment questions about their voices, which required a reply, and then dependent on their responses, a thought-challenging message for unhelpful beliefs and a behavioural coping strategy or experiment suggestion.

### *Feasibility*

Overall, the technology assisted psychological assessments and treatments were considered feasible to implement; researchers were able to administer the intended assessment or treatment, participants could use the technologies with minimal technical difficulties and there were generally low study dropouts and withdrawals. There were no reported withdrawals or dropouts from the assessment studies, which assessed voices using computerised avatar software (Cardi *et al.*, 2022) or using VR (Stinson *et al.*, 2010). In the treatment studies, avatar therapy using computerised software was reported as feasible to deliver (Leff *et al.*, 2013; Craig *et al.*, 2018; Rus-Calafell *et al.*, 2020). Initially, there were very few referrals for the therapy in the early pilot trial by Leff *et al.* (2013) but a steady rate of referrals was achieved after some successes with the therapy. A high drop-out rate was reported in this pilot study, with around a third of participants dropping out. In the large RCT that followed, the retention rate of participants at the 24 week follow-up was 76% (Craig *et al.*, 2018). Avatar-based therapy using immersive VR appeared feasible to deliver, with retention rates above 75%. Findings from the studies on the computerised, Web-based CBTp programme, mobile device-assisted treatment and the apps, suggest they were feasible to deliver, with high retention rates and most mobile phones (86%) returned intact for the mobile-based treatment. However, in the large naturalistic cohort study on the Netherlands app, most app-users (71%) were excluded due to not reaching level two or playing less than 15 times.

### *Acceptability*

Overall, studies indicate that people with distressing voices found the technology acceptable for assessment and treatment. There were high levels of satisfaction reported by participants who used the technologies. When computerised avatar software was used to assess the eating disorder voice, all participants said they would be willing to be re-exposed to the avatar in the future, although almost 90% of individuals reported some level of

distress in response to exposure to the computerised 2D avatar (Cardi *et al.*, 2022). When distressing voices were assessed using VR, Stinson *et al.* (2010) reported no differences in participants' anxiety levels and heart rates before and after VR and a majority (93%) of participants reported that they did not experience unwanted thoughts about the VR in the week following the study.

The avatar-based treatment studies reported mostly positive feedback from participants. When conventional, computerised software was used, there were no adverse effects attributed to the therapy, with reasons for discontinuation including logistical issues, physical health problems or participants reporting that the approach was not relevant or helpful for them (Craig *et al.*, 2018). Researchers reported that facing the avatar took considerable courage for some participants, and two participants heard multiple voices, so they could not concentrate on the avatar because the other voices spoke too loudly at the same time. In the avatar-based studies that used *immersive* VR, participants reported that the immersive environment helped to make their experience "come to life" by enabling a more direct discussion with their voice (e.g. "the avatar was truly there"). Although almost 40% reported that the treatment was stressful at first, they reported that they enjoyed their experience (Du Sert *et al.*, 2018; Dellazizzo *et al.*, 2020, 2021).

Participants who engaged in the Web-based CBTp programme, mobile-device assisted treatment and app-based treatments also reported positive feedback. Most of the participants thought the Web-based and mobile phone-based treatments were "helpful" or "very helpful" (Granholm *et al.*, 2012; Gottlieb *et al.*, 2013, 2017). User feedback on the apps included that they liked that the app was consistently available to them and that they were able to access helpful tools in the moment (Bell *et al.*, 2020; Buck *et al.*, 2022). There was no participant feedback from users of the Temstem app in the large naturalistic the Netherlands study (Jongeneel *et al.*, 2022).

### *Effectiveness*

Overall, the digital technologies were effective in assessing distressing voices, as well as effectively reducing the frequency and severity of distressing voices and reducing voice-related distress. The Cardi *et al.* (2022) assessment study on the eating disorder voice using computerised avatar software indicated that there was a "very good" or "good" match between the sound of the avatar and the imagined sound of their internal voice. In the VR assessment study, participants experienced voices in VR, but these were not found to be triggered by cognitive antecedents (Stinson *et al.*, 2010).

Avatar therapy, with both computerised software on 2D screens and immersive VR, was effective in reducing voice frequency, severity, omnipotence and associated distress. However, there was no significant difference reported between the avatar group and supportive counselling at 24-week follow-up in the larger clinical trial (Craig *et al.*, 2018). The "Coping with Voices" computerised CBTp programme was associated with significant voice-related improvements in the 2013 pilot study, including overall severity, perception of voices as an outside entity, negative commentary from voices and trends for reduced frequency and durations of voices, reduced perception that voices emanated from an external source and increased perceived control over voices (Gottlieb *et al.*, 2013). The mobile-assisted treatment was associated with a significant reduction in being bothered by voices over the course of treatment (Granholm *et al.*, 2012). The three mobile app studies had promising results on their effectiveness, with two apps associated with significant reduction in voice frequency, severity and voice-related distress (Bell *et al.*, 2020; Jongeneel *et al.*, 2022). One app was associated with small positive effects on severity of voices, although these improvements were not statistically significant (Buck *et al.*, 2022).

### Quality ratings

Overall, EPHPP ratings were “strong” ( $n = 8$ ), “moderate” ( $n = 5$ ) and “weak” ( $n = 1$ ). The EPHPP global ratings for the two assessment studies were “strong” and “weak”. The treatment studies received global ratings of “strong” ( $n = 7$ ) and “moderate” ( $n = 5$ ). The study design and data collection method domains were generally rated as “strong”, while the selection bias and blinding domains were rated as “moderate” for all studies. See [Table 2](#) for full details of the quality assessment.

### Discussion

The aim of this systematic review was to evaluate technology-enhanced psychological assessment and treatment for distressing voices. Overall, studies indicated that technologies are feasible to deliver and acceptable to users with distressing voices. Preliminary evidence suggests that the technologies are effective in reducing the frequency and severity of distressing voices, but there is still a relatively small number of studies in this area, with more RCTs needed to establish effectiveness. The fact that over half of the studies were published in the past five years reflects the significant recent developments in this area, particularly avatar therapy. All studies, except for one, sampled clinical groups, mostly consisting of those with a diagnosis of a schizophrenia spectrum or psychotic disorder, reflecting the extensive investigation of distressing voices in the context of these conditions. Despite increasing interest in the literature on the “eating disorder voice”, often experienced by those diagnosed with an eating disorder as a voice, which is powerful, negative and omnipotent in nature ([Aya et al., 2019](#)), only one study in the review sampled this group.

Most studies in the review investigated distressing voices using avatar-based approaches, using either computerised software on 2D screens ([Leff et al., 2013](#); [Craig et al., 2018](#)) or *immersive* VR ([Du Sert et al., 2018](#); [Dellazizzo et al., 2020, 2021](#)). Avatar therapy offers a unique opportunity for voice-hearers to have a direct dialogue with a digital representation of their voice (the avatar) and challenge their threat-based beliefs about their voices, and there has been rapidly growing interest in this novel, relational approach to working with distressing voices. Evidence suggests that avatar therapy is feasible to deliver and, despite some initial concerns from participants in the early trial, overall, positive user experience has been reported. Initial concerns from participants reflected in high dropout rates in the pilot trial are understandable and somewhat expected, given the untried nature of the therapy at that stage. Avatar therapy shows great promise as an effective treatment for distressing voices, as it is associated with significant reductions in the frequency, severity and omnipotence of the voice, as well as voice-related distress. Promising findings on this innovative treatment are supported by the thematic qualitative evaluation, which was part of the large [Craig et al. \(2018\)](#) trial and involved semi-structured interviews with participants in the avatar therapy group ([Rus-Calafell et al., 2022](#)). Participants reported an overall positive experience and described voice-related improvements affecting their everyday life outside of therapy. Notably, all of those interviewed highlighted the relationship with the therapist and reported feeling supported and understood by the therapist throughout the therapy sessions, reflecting the importance of in-person, therapeutic support when administering avatar therapy.

The highly interactive computerised CBTp programme and the mobile-assisted psychological treatment seemed feasible to deliver and acceptable to participants with distressing voices ([Granholtm et al., 2012](#); [Gottlieb et al., 2013; 2017](#)). There were high study completion rates, and most phones were returned undamaged at the end of the study on the mobile-assisted treatment. Most participants reported that they would recommend the treatment to a friend or relative. The treatments were associated with significantly reduced



**Table 2.**  
Summary of effective  
public health practice  
project quality  
ratings of the  
included studies

Study	Selection bias	Study design	Confounders	Blinding	Data collection method	Withdrawal and drop-out	Global rating
<i>Assessment studies</i>							
Stinson <i>et al.</i> (2010)	Moderate	Strong	Strong	Moderate	Strong	Strong	Strong
Cardi <i>et al.</i> (2022)	Moderate	Moderate	Weak	Moderate	Weak	N/A	Weak
<i>Treatment studies</i>							
Bell <i>et al.</i> (2020)	Moderate	Strong	Strong	Moderate	Strong	Strong	Strong
Buck <i>et al.</i> (2022)	Moderate	Moderate	N/A	Moderate	Strong	N/A	Strong
Craig <i>et al.</i> (2018)	Moderate	Strong	Weak	Moderate	Strong	Moderate	Moderate
Dellazizzo <i>et al.</i> (2020)	Moderate	Moderate	N/A	Moderate	Strong	N/A	Strong
Dellazizzo <i>et al.</i> (2021)	Moderate	Strong	Strong	Moderate	Strong	Weak	Moderate
Du Sert <i>et al.</i> (2018)	Moderate	Strong	Strong	Moderate	Strong	Moderate	Strong
Gottlieb <i>et al.</i> (2013)	Moderate	Moderate	N/A	Moderate	Strong	Strong	Strong
Gottlieb <i>et al.</i> (2017)	Moderate	Strong	Weak	Moderate	Strong	Strong	Moderate
Granhölm <i>et al.</i> (2012)	Moderate	Moderate	N/A	Moderate	Weak	Moderate	Moderate
Jongeneel <i>et al.</i> (2022)	Moderate	Moderate	N/A	Moderate	Weak	N/A	Moderate
Leff <i>et al.</i> (2013)	Moderate	Strong	Strong	Moderate	Strong	Moderate	Strong
Rus-Calafell <i>et al.</i> (2020)	Moderate	Moderate	N/A	Moderate	Strong	Strong	Strong

**Source:** Created by authors

frequency and severity of voices, although voice-related improvements associated with the Web-based CBTp programme were comparable to those seen in the control group in the [Gottlieb et al. \(2017\)](#) trial. Significant improvements in social functioning and knowledge in CBTp following the programme offer some support for the technology-enhanced treatment functioning as a steppingstone to in-person therapy ([Bond et al., 2023](#)).

There were only three studies in the review that used mobile apps to treat distressing voices, despite the exponential growth in the use of mobile phones, alongside mobile apps, in the past decade ([Laricchia, 2023](#)). Rates of smartphone ownership and use of smartphones by people with serious mental illness have been reported as similar to those of the general population ([Young et al., 2020](#)) and findings from the current review highlight apps effectively reducing the frequency of distressing voices ([Bell et al., 2020](#); [Buck et al., 2022](#); [Jongeneel et al., 2022](#)). Generally, they appeared to be feasible and acceptable, although one large, naturalistic study investigating the “Temstem” app had low completion rates ([Jongeneel et al., 2022](#)). This is consistent with real-world objective data on app user engagement, highlighting that many users of mental health apps simply download and delete the app without using the app for any sustained period ([Baumel et al., 2019](#)). However, high completion rates have been reported for apps targeting psychotic symptoms in the context of controlled research studies, whereby participants were provided with smartphones and/or internet access ([Clarke et al., 2019](#)). Evidently, apps might be a promising tool for treating distressing voices, if they can be sufficiently engaging, but more research is needed in this area and a combination of controlled trials and naturalistic studies would be helpful.

#### *Strengths and limitations of the literature*

A strength of the literature is that a majority (13 out of 14) of studies received a global quality rating of “strong” or “moderate”, indicating the high-quality research emerging in this area, despite there still being a relatively small number of studies published. Further, all included study designs that were rated as “strong” or “moderate”, and reliable and valid measures of voices were consistently used in the literature.

However, there were few studies with a large and representative sample and there were only a few RCTs. Additionally, a limitation is that, although studies consistently reported participants’ age and gender, five of the studies did not report participants’ ethnicity. Studies that reported participants’ ethnicity reported a majority of participants from a White ethnic background, including one study with all participants identifying as Caucasian ([Dellazizzo et al., 2020](#)). This lack of reporting and lack of ethnic diversity reported make it difficult to generalise the findings to voice-hearers from a diverse range of ethnic backgrounds.

#### *Strengths and limitations of the review*

A strength of the review is that it is the first to investigate technology-enhanced assessment and treatment of distressing voices, using a transdiagnostic and continuum approach to voices. Encompassing a wide spectrum of experiences aligns with the evolving understanding of voice hearing. This approach acknowledges the diverse nature of voice hearing experiences, contributing to a more comprehensive understanding of assessment and treatment for voices. The symptom-specific approach is another strength of the current review, as most previous studies on digital health interventions have targeted psychosis, and there have been calls for future psychosis research to develop a symptom-specific approach ([Clarke et al., 2019](#)). However, a potential implication of this specific approach to the review is potentially overlooking studies on digital health interventions that impact

voices indirectly or were not explicitly designed to target voices. If studies with broader focuses did not explicitly mention voices, they may have been excluded from the review, potentially limiting the scope of the findings.

The focus on technologies with an “enhancing” function in the current review presents a specific study inclusion criterion, emphasising the role of technology in augmenting psychological assessment or treatment. However, this distinction introduces subjectivity and potential challenges, particularly in determining borderline cases, which may also have led to studies being excluded from the review. Additionally, the subjective element in the use of the EPHPP quality rating tool indicates that the assessment of study quality involves a certain degree of interpretation.

### *Clinical implications*

There is a range of potential opportunities, as well as challenges, associated with integrating technology-enhanced assessments and treatments into clinical practice.

Promising findings from the review indicate that technologies can potentially function as valuable tools to support existing mental health services, enhancing the delivery of psychological care. In the context of stretched NHS mental health services and lengthy waiting times, technology can increase access and engagement in psychological assessments and treatments. Particularly, computerised and app-based treatments offer remote accessibility, addressing geographical, financial and logistic barriers to accessing mental health services (Torous *et al.*, 2021).

Avatar therapy provides a novel and exciting approach to working with distressing voices, empowering individuals to actively engage with and confront their voices. However, challenges include the high cost of equipment, which may be inaccessible to many and the need for specially trained therapists, acknowledged by the developers as an issue for rolling this therapy out across clinical settings in a cost-effectiveness manner.

Further, the concept of a “digital divide”, where certain populations, particularly those with psychosis, may face barriers to technology access is crucial to acknowledge and address to ensure equitable access to technology-enhanced assessments and treatments (Watson *et al.*, 2022). Staff in secondary care mental health teams have identified the “digital divide” as a significant potential barrier to implementing digital tools in clinical practice (Bucci *et al.*, 2019). Groups identified as particularly impacted by this are those who are older, have persistent psychotic conditions (Young *et al.*, 2020) and ethnic minority groups (Aref-Adib *et al.*, 2019).

Despite the promising opportunities, there are recognised challenges in integrating and implementing these technologies into routine clinical practice. Factors such as immediate costs, lack of IT support, infrastructure limitations and the need for involvement of frontline staff and end-users in design and rollout are critical considerations (Aref-Adib *et al.*, 2019). Balancing innovation with practical considerations and addressing barriers to accessibility are key considerations for supporting successful integration of these technologies into mental health care.

### *Future directions*

The review included peer-reviewed, empirical studies and, therefore, only high-quality research was included. However, there was a relatively small number of included studies in this review, indicating a need for further, high-quality research in this area. Several studies are underway on technology-enhanced treatments for distressing voices and there will likely be significant developments in this area in the next 5–10 years. Of note, a study protocol for the AVATAR2 RCT has been published to further explore the efficacy of this novel therapy

(Garety *et al.*, 2021), as well as a study protocol for a large RCT on a novel smartphone-assisted guided CBT intervention for distressing voices (Cavelti *et al.*, 2022).

Additionally, the current review highlighted the lack of reporting on participant ethnicity data, as well as a lack of ethnic diversity in samples in the literature, which is a critical observation. Future studies should prioritise ethnic diversity in sampling and consider cultural adaptations, which aligns with broader calls for increased diversity in digital mental health research (Jiménez-Molina *et al.*, 2019; Riches *et al.*, 2023).

The relatively limited range of technologies specifically targeting distressing voices points to a potential avenue for future research. For instance, the review did not include any studies on wearables to passively monitor symptoms, despite evidence that individuals with psychosis are comfortable, able and willing to use these devices (Cella *et al.*, 2019). Exploring a broader range of technologies, including wearables and artificial intelligence, will be a useful direction for future research.

Overall, there are promising findings from this review, but more high-quality research is warranted to establish the feasibility, acceptability and effectiveness of the technology-enhanced assessments and treatments. A focus on ethnic diversity, exploration of varied technologies, qualitative insights and ethical considerations, would be useful.

### Conclusion

Overall, technology-enhanced psychological assessments and treatments for distressing voices appear feasible, acceptable and effective, indicating the potential for these technologies to enhance mental health care. Avatar therapy has been the most researched treatment, but more research is needed to refine and advance this innovative therapeutic approach.

Collaborative efforts between researchers, clinicians and individuals with distressing voices are crucial for the ongoing development of technology-enhanced assessments and treatments. Integration and implementation into clinical practice is a recognised challenge, although the findings from the review offers some promising evidence for the clinical benefits of these technologies enhancing mental health care for individuals with distressing voices.

### References

- Aref-Adib, G., McCloud, T., Ross, J., O'Hanlon, P., Appleton, V., Rowe, S., Murray, E., Johnson, S. and Lobban, F. (2019), "Factors affecting implementation of digital health interventions for people with psychosis or bipolar disorder, and their family and friends: a systematic review", *The Lancet Psychiatry*, Vol. 6 No. 3, pp. 257-266.
- Aya, V., Ulusoy, K. and Cardi, V. (2019), "A systematic review of the 'eating disorder voice' experience", *International Review of Psychiatry*, Vol. 31 No. 4, pp. 347-366.
- Baños, R.M., Herrero, R. and Vara, M.D. (2022), "What is the current and future status of digital mental health interventions?", *The Spanish Journal of Psychology*, Vol. 25, p. e5.
- Baumeister, D., Sedgwick, O., Howes, O. and Peters, E. (2017), "Auditory verbal hallucinations and continuum models of psychosis: a systematic review of the healthy voice-hearer literature", *Clinical Psychology Review*, Vol. 51, pp. 125-141.
- Baumel, A., Muench, F., Edan, S. and Kane, J.M. (2019), "Objective user engagement with mental health apps: systematic search and panel-based usage analysis", *Journal of Medical Internet Research*, Vol. 21 No. 9, p. e14567.
- Bell, I.H., Lim, M.H., Rossell, S.L. and Thomas, N. (2017), "Ecological momentary assessment and intervention in the treatment of psychotic disorders: a systematic review", *Psychiatric Services*, Vol. 68 No. 11, pp. 1172-1181.

- Bell, I.H., Rossell, S.L., Farhall, J., Hayward, M., Lim, M.H., Fielding-Smith, S.F. and Thomas, N. (2020), "Pilot randomised controlled trial of a brief coping-focused intervention for hearing voices blended with smartphone-based ecological momentary assessment and intervention (SAVVy): feasibility, acceptability and preliminary clinical outcomes", *Schizophrenia Research*, Vol. 216, pp. 479-487.
- Bond, R.R., Mulvenna, M.D., Potts, C., O'Neill, S., Ennis, E. and Torous, J. (2023), "Digital transformation of mental health services", *Npj Mental Health Research*, Vol. 2 No. 1, p. 13.
- Bonet, L., Izquierdo, C., Escarti, M.J., Sancho, J.V., Arce, D., Blanquer, I. and Sanjuan, J. (2017), "Use of mobile technologies in patients with psychosis: a systematic review", *Revista de Psiquiatria y Salud Mental (English Edition)*, Vol. 10 No. 3, pp. 168-178.
- Bucci, S., Barrowclough, C., Ainsworth, J., Machin, M., Morris, R., Berry, K., Emsley, R., Lewis, S., Edge, D., Buchan, I. and Haddock, G. (2018), "Actissist: proof-of-concept trial of a theory-driven digital intervention for psychosis", *Schizophrenia Bulletin*, Vol. 44 No. 5, pp. 1070-1080.
- Bucci, S., Berry, N., Morris, R., Berry, K., Haddock, G., Lewis, S. and Edge, D. (2019), "They are not hard-to-reach clients. We have just got hard-to-reach services." staff views of digital health tools in specialist mental health services", *Frontiers in Psychiatry*, Vol. 10, p. 344.
- Buck, B., Nguyen, J., Porter, S., Ben-Zeev, D. and Reger, G.M. (2022), "FOCUS mHealth intervention for veterans with serious mental illness in an outpatient department of veterans' affairs setting: feasibility, acceptability, and usability study", *JMIR Mental Health*, Vol. 9 No. 1, p. e26049.
- Campbell, M., McKenzie, J.E., Sowden, A., Katikireddi, S.V., Brennan, S.E., Ellis, S., Hartmann-Boyce, J., Ryan, R., Shepperd, S., Thomas, J. and Welch, V. (2020), "Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline", *BMJ*, p. 368.
- Cardi, V., Ward, T., Aya, V., Calissano, C., Thompson, A. and Treasure, J. (2022), "A proof-of-concept study for the use of a computerised avatar to embody the eating disorder voice in anorexia nervosa", *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity*, Vol. 27 No. 8, pp. 3499-3506.
- Cavelti, M., Kaeser, J.M., Lerch, S., Bauer, S., Moessner, M., Berger, T., Hayward, M. and Kaess, M. (2022), "Smartphone-assisted guided self-help cognitive behavioral therapy for young people with distressing voices (SmartVoices): study protocol for a randomized controlled trial", *Trials*, Vol. 23 No. 1, p. 902.
- Cella, M., He, Z., Killikelly, C., Okruszek, L., Lewis, S. and Wykes, T. (2019), "Blending active and passive digital technology methods to improve symptom monitoring in early psychosis", *Early Intervention in Psychiatry*, Vol. 13 No. 5, pp. 1271-1275.
- Ciliska, D., Miccouci, S. and Dobbins, M. (1998), *Effective Public Health Practice Project. quality Assessment Tool for Quantitative Studies*, Effective Public Health Practice Project, Hamilton, On.
- Clarke, S., Hanna, D., Mulholland, C., Shannon, C. and Urquhart, C. (2019), "A systematic review and meta-analysis of digital health technologies effects on psychotic symptoms in adults with psychosis", *Psychosis*, Vol. 11 No. 4, pp. 362-373.
- Connell, M., Scott, J.G., McGrath, J.J., Waters, F., Laroi, F., Alati, R., Najman, J. and Betts, K. (2019), "A comparison of hallucinatory experiences and their appraisals in those with and without mental illness", *Psychiatry Research*, Vol. 274, pp. 294-300.
- Craig, T.K., Rus-Calafell, M., Ward, T., Leff, J.P., Huckvale, M., Howarth, E., Emsley, R. and Garety, P.A. (2018), "AVATAR therapy for auditory verbal hallucinations in people with psychosis: a single-blind, randomised controlled trial", *The Lancet Psychiatry*, Vol. 5 No. 1, pp. 31-40.
- Dellazizzo, L., Potvin, S., Phraxayavong, K. and Dumais, A. (2020), "Exploring the benefits of virtual reality-assisted therapy following cognitive-behavioral therapy for auditory hallucinations in patients with treatment-resistant schizophrenia: a proof of concept", *Journal of Clinical Medicine*, Vol. 9 No. 10, p. 3169.
- Dellazizzo, L., Potvin, S., Phraxayavong, K. and Dumais, A. (2021), "One-year randomized trial comparing virtual reality-assisted therapy to cognitive-behavioral therapy for patients with treatment-resistant schizophrenia", *Npj Schizophrenia*, Vol. 7 No. 1, p. 9.

- Du Sert, O.P., Potvin, S., Lipp, O., Dellazizzo, L., Laurelli, M., Breton, R., Lalonde, P., Phraxayavong, K., O'Connor, K., Pelletier, J.F. and Boukhalfi, T. (2018), "Virtual reality therapy for refractory auditory verbal hallucinations in schizophrenia: a pilot clinical trial", *Schizophrenia Research*, Vol. 197, pp. 176-181.
- Garety, P., Edwards, C.J., Ward, T., Emsley, R., Huckvale, M., McCrone, P., Rus-Calafell, M., Fornells-Ambrojo, M., Gumley, A., Haddock, G. and Bucci, S. (2021), "Optimising AVATAR therapy for people who hear distressing voices: study protocol for the AVATAR2 multi-Centre randomised controlled trial", *Trials*, Vol. 22 No. 1, pp. 1-17.
- Gottlieb, J.D., Romeo, K.H., Penn, D.L., Mueser, K.T. and Chiko, B.P. (2013), "Web-based cognitive-behavioral therapy for auditory hallucinations in persons with psychosis: a pilot study", *Schizophrenia Research*, Vol. 145 No. 1-3, pp. 82-87.
- Gottlieb, J.D., Gidugu, V., Maru, M., Tepper, M.C., Davis, M.J., Greenwold, J., Barron, R.A., Chiko, B.P. and Mueser, K.T. (2017), "Randomized controlled trial of an internet cognitive behavioral skills-based program for auditory hallucinations in persons with psychosis", *Psychiatric Rehabilitation Journal*, Vol. 40 No. 3, p. 283.
- Granhölm, E., Ben-Zeev, D., Link, P.C., Bradshaw, K.R. and Holden, J.L. (2012), "Mobile assessment and treatment for schizophrenia (MATS): a pilot trial of an interactive text-messaging intervention for medication adherence, socialization, and auditory hallucinations", *Schizophrenia Bulletin*, Vol. 38 No. 3, pp. 414-425.
- Jauhar, S., McKenna, P.J., Radua, J., Fung, E., Salvador, R. and Laws, K.R. (2014), "Cognitive-behavioural therapy for the symptoms of schizophrenia: systematic review and meta-analysis with examination of potential bias", *British Journal of Psychiatry*, Vol. 204 No. 1, pp. 20-29.
- Jiménez-Molina, Á., Franco, P., Martínez, V., Martínez, P., Rojas, G. and Araya, R. (2019), "Internet-based interventions for the prevention and treatment of mental disorders in latin america: a scoping review", *Frontiers in Psychiatry*, Vol. 10, p. 664.
- Jongeneel, A., Libedinsky, I., Reinbergen, A., Tromp, N., Delespaul, P., Riper, H., van der Gaag, M. and van den Berg, D. (2022), "Momentary effects of temstem, an app for voice-hearing individuals: Results from naturalistic data from 1048 users", *Internet Interventions*, Vol. 30, p. 100580.
- Laricchia, F. (2023), "Smartphone ownership in the United Kingdom (UK) 2012-2023, by age", available at: [www.statista.com/statistics/271851/smartphone-owners-in-the-united-kingdom-uk-by-age/](http://www.statista.com/statistics/271851/smartphone-owners-in-the-united-kingdom-uk-by-age/) (accessed 4 February 2024).
- Leff, J., Williams, G., Huckvale, M.A., Arbutnot, M. and Leff, A.P. (2013), "Computer-assisted therapy for medication-resistant auditory hallucinations: proof-of-concept study", *British Journal of Psychiatry*, Vol. 202 No. 6, pp. 428-433.
- Maijjer, K., Begemann, M.J., Palmen, S.J., Leucht, S.J.M.C. and Sommer, I.E. (2018), "Auditory hallucinations across the lifespan: a systematic review and meta-analysis", *Psychological Medicine*, Vol. 48 No. 6, pp. 879-888.
- Maroño Souto, Y., Vázquez Campo, M., Díaz Llenderozas, F., Rodríguez Álvarez, M., Mateos, R. and García Caballero, A. (2018), "Randomized clinical trial with e-MotionalTraining® 1.0 for social cognition rehabilitation in schizophrenia", "p", *Frontiers in Psychiatry*, Vol. 9, p. 40.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., Prisma Group. \*, (2009), "Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement", *Annals of Internal Medicine*, Vol. 151 No. 4, pp. 264-269.
- Riches, S., Fallah, D. and Kaleva, I. (2023), "Increasing diversity and inclusion in research on virtual reality relaxation: Commentary on 'virtual reality relaxation for people with mental health conditions: a systematic review'", *Journal of Mental Health and Clinical Psychology*, Vol. 7 No. 2, pp. 24-29.
- Riches, S., Pisani, S., Bird, L., Rus-Calafell, M., Garety, P. and Valmaggia, L. (2021), "Virtual reality-based assessment and treatment of social functioning impairments in psychosis: a systematic review", *International Review of Psychiatry*, Vol. 33 No. 3, pp. 337-362.



- Rus-Calafell, M., Ehrbar, N., Ward, T., Edwards, C., Huckvale, M., Walke, J., Garety, P. and Craig, T. (2022), "Participants' experiences of AVATAR therapy for distressing voices: a thematic qualitative evaluation", *BMC Psychiatry*, Vol. 22 No. 1, p. 356.
- Rus-Calafell, M., Ward, T., Zhang, X.C., Edwards, C.J., Garety, P. and Craig, T. (2020), "The role of sense of voice presence and anxiety reduction in AVATAR therapy", *Journal of Clinical Medicine*, Vol. 9 No. 9, p. 2748.
- Schutte, M.J., Linszen, M.M., Marschall, T.M., Koops, S., van Dellen, E., Heringa, S.M., Slooter, A.J., Teunisse, R., van den Heuvel, O.A., Lemstra, A.W. and Foncke, E.M. (2020), "Hallucinations and other psychotic experiences across diagnoses: a comparison of phenomenological features", *Psychiatry Research*, Vol. 292, p. 113314.
- Steinhubl, S.R., Muse, E.D. and Topol, E.J. (2013), "Can mobile health technologies transform health care?", *JAMA*, Vol. 310 No. 22, pp. 2395-2396.
- Stinson, K., Valmaggia, L.R., Antley, A., Slater, M. and Freeman, D. (2010), "Cognitive triggers of auditory hallucinations: an experimental investigation", *Journal of Behavior Therapy and Experimental Psychiatry*, Vol. 41 No. 3, pp. 179-184.
- Thomas, B.H., Ciliska, D., Dobbins, M. and Micucci, S. (2004), "A process for systematically reviewing the literature: providing the research evidence for public health nursing interventions", *Worldviews on Evidence-Based Nursing*, Vol. 1 No. 3, pp. 176-184.
- Toh, W.L., Moseley, P. and Fernyhough, C. (2022), "Hearing voices as a feature of typical and psychopathological experience", *Nature Reviews Psychology*, Vol. 1 No. 2, pp. 72-86.
- Torous, J., Bucci, S., Bell, I.H., Kessing, L.V., Faurholt-Jepsen, M., Whelan, P., Carvalho, A.F., Keshavan, M., Linardon, J. and Firth, J. (2021), "The growing field of digital psychiatry: current evidence and the future of apps, social media, chatbots, and virtual reality", *World Psychiatry*, Vol. 20 No. 3, pp. 318-335.
- Turner, D.T., van der Gaag, M., Karyotaki, E. and Cuijpers, P. (2014), "Psychological interventions for psychosis: a meta-analysis of comparative outcome studies", *American Journal of Psychiatry*, Vol. 171 No. 5, pp. 523-538.
- Watson, A., Mellotte, H., Hardy, A., Peters, E., Keen, N. and Kane, F. (2022), "The digital divide: factors impacting on uptake of remote therapy in a South London psychological therapy service for people with psychosis", *Journal of Mental Health*, Vol. 31 No. 6, pp. 825-832.
- Wykes, T., Steel, C., Everitt, B. and Tarriner, N. (2008), "Cognitive behavior therapy for schizophrenia: effect sizes, clinical models, and methodological rigor", *Schizophrenia Bulletin*, Vol. 34 No. 3, pp. 523-537.
- Young, A.S., Cohen, A.N., Niv, N., Nowlin-Finch, N., Oberman, R.S., Olmos-Ochoa, T.T., Goldberg, R.W. and Whelan, F. (2020), "Mobile phone and smartphone use by people with serious mental illness", *Psychiatric Services*, Vol. 71 No. 3, p. 28.

**Corresponding author**

Simon Riches can be contacted at: [simon.j.riches@kcl.ac.uk](mailto:simon.j.riches@kcl.ac.uk)