University of Glasgow

Caird, Julia S.G. (2023) *Trauma-hallucination phenomenology and the psychometric properties of Visual Hallucination measures.* D Clin Psy thesis.

http://theses.gla.ac.uk/83998/

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten: Theses
https://theses.gla.ac.uk/
research-enlighten@glasgow.ac.uk

Trauma-hallucination phenomenology and the psychometric properties of Visual Hallucination measures

Julia S.G Caird, MA (Hons), PGDip

Submitted in partial fulfilment of the requirements for the degree of Doctorate in Clinical Psychology

School of Health and Wellbeing
College of Medical, Veterinary and Life Sciences
University of Glasgow

December 2023

Table of Contents

List of Tables 3
List of Figures 4
Acknowledgement 5
Chapter 1: The Reliability and Validity of current measures of visual hallucinations in adult populations: A systematic review 7
Abstract 8
Introduction 9
Methods 12
Results 17
Discussion 36
References 42
Chapter 2: Exploring the phenomenology of hallucinatory experiences and their relation to trauma re-experiencing symptoms: The Questionnaire of Hallucinatory Experiences (QHE) 47
Plain Language Summary 48
Abstract 50
Introduction 51
Methods 55
Results 62
Discussion 83
References 89
Appendix 1.1: PRISMA 2020 Checklist 98
Appendix 1.2: Search Strategy 102
Appendix 1.3: Overview of Measures included in the systematic review. 105
Appendix 1.4: Updated COSMIN criteria for good measurement properties .. 111
Appendix 1.5: Measures excluded from Aynsworth et al., (2017) review. 113
Appendix 1.6: COSMIN Criteria for sufficient content validity 114
Appendix 1.7: Overall Content Validity and Quality of Evidence Ratings 115
Appendix 1.8: Psychometric properties reported in the included articles. 116
Appendix 1.9 Overall Psychometric Quality Ratings and Rationale 119
Appendix 2.1: MRP Proposal 121
Appendix 2.2 Revised Analysis Plan 122
Appendix 2.3 Ethical Approval 123
Appendix 2.4 Questionnaire of Hallucinatory Experiences (QHE) 124
Appendix 2.5 Consent Form 125
Appendix 2.6 Participant Information Sheet 126

List of Tables

Chapter 1

Table 1. 1: COSMIN definitions of measurement properties 15
Table 1. 2: Overview of included articles 21
Table 1. 3: Quality of Measure Development. 28
Table 1. 4: Psychometric quality and overall evidence quality ratings as defined by the GRADE approach. 32

Chapter 2

Table 2.1: Scales and assessment measures reviewed during development of the QHE. 57
Table 2.2: Go/Revise/Stop Criteria for data analysis 61
Table 2.3: Demographic Characteristics of Participants 62
Table 2.4: Subscale Internal Consistency–Auditory Hallucination (AH) Phenomenology Scale 64
Table 2.5: Subscale Internal Consistency – Visual Hallucination (VH) Phenomenology Scale 65
Table 2.6: Subscale Internal Consistency-Tactile Hallucination (TH) Phenomenology Scale 66
Table 2.7: Subscale Internal Consistency- Sensed Presence (SP) Hallucination Phenomenology Subscale 67
Table 2.8: Subscale Internal Consistency- Multi-Modal (MMH) Hallucination Phenomenology Subscale 68
Table 2.9: Subscale Internal Consistency – Auditory Hallucination (AH) Relation to Traumatic Experiences Scale 70
Table 2.10: Subscale Internal Consistency – Visual Hallucination (VH) Relation to Traumatic Experience Scale 71
Table 2.11: Subscale Internal Consistency-Tactile Hallucination (TH) Relation to Traumatic Experiences Scale 72
Table 2.12: Subscale Internal Consistency- Sensed Presence (SP) Hallucination Relation to Traumatic Experiences Subscale 74
Table 2.13: Subscale Internal Consistency- Multi-Modal (MMH) Hallucination Relation to Traumatic Experiences Subscale 75
Table 2.14: Associations between hallucinatory phenomenology and trauma subscales 78
Table 2.15: Summary of Analyses According to Go/Revise/Stop Criteria 81

List of Figures

Chapter 1

Figure 1.1: PRISMA flow diagram for identification of study 20

Chapter 2

Figure 2.1: Study Design Phases 55
Figure 2.2: QHE modules and structure..... 59
Figure 2.3: Number of participants reporting the experience of hallucinations (by modality) in the past month. 63

Acknowledgement

Firstly, I would like to thank the participants of the Major Research Project. Without you, this project would not have been possible. My heartfelt thanks for your contribution, time, and openness.

A sincere thank you to Professor Andrew Gumley for your dynamic thinking, encouragement, support, and patience, especially towards the end of the project when it seemed there was still a mountain to climb. A massive thank you to my field supervisor, Dr. Kirsten Atherton, whose compassion and containing manner pulled me through, as did her skills and persistence with promoting the study on Twitter. Thank you to my field supervisor Dr. Andy Siddaway, the scale development guru, for his guidance, pragmatism, and expertise.

I am eternally grateful to Professor Filipo Varese, Dr. Eleanor Longden, Dr. Amy Hardy, Dr. Moya Clancy, Dr. Stephanie Allan, Dr. Samantha Bowe, Dr. Leonie Richardson and Professor Ben Alderson-Day for your invaluable feedback on the initial QHE items and advice for revision. Thank you to Dr. Rufus May for his help with recruitment and Professor Nev Jones for her suggestions for further QHE iteration. I would also like to thank Jenny McGhie for her patience and kindness in answering my many systematic review queries, and Paul Cannon for his assistance in developing a search strategy.

A massive thanks to my third-year placement supervisors, Dr. Jen Megson-Gray, Dr. Jane Burns and Dr. Lilian Wanless for their unwavering encouragement, cheerleading, and wisdom. To Jenny, my second rater, fellow mum trainee and friend, thank you for always being there. We did it, and with toddlers!

To my mum, a special thank you for your constant belief in me and late-night proof reading. A heartfelt thanks to Sarah for her patience, unwavering support, and formatting

skills. To my dear friends from the course, thank you for your ever-present understanding and humour, especially during the pandemic when we couldn't be together. Lastly, to Mark who I can't thank enough, for being with me every step of the way while I pursued a career in Clinical Psychology. You have been largely lone parenting for the last year and a half. Your tireless support, commitment, and ability to look after Alfie and I means the world; I could not have done this without you.

Chapter 1: The Reliability and Validity of current measures of visual hallucinations in adult populations: A systematic review

Authors: Julia Caird, Jenny McKeever, Professor Andrew Gumley

1. Mental Health and Wellbeing, School of Health and Wellbeing, University of Glasgow; and NHS Greater Glasgow and Clyde

2. Mental Health and Wellbeing, School of Health and Wellbeing, University of Glasgow; and NHS Greater Glasgow and Clyde

3. Mental Health and Wellbeing, School of Health and Wellbeing, University of Glasgow

Corresponding address:

Clarice Pears Building
University of Glasgow
90 Byres Road
Glasgow
G12 8TB
j.caird.1@research.gla.ac.uk

Prepared in accordance with Clinical Psychology Review: Guide for authors - Clinical Psychology Review - ISSN 0272-7358 | ScienceDirect.com by Elsevier

Word Count: 8,540

Abstract

Background: Frequently distressing, visual hallucinations (VH) are found in a range of organic and psychiatric disorders and in non-clinical populations. Reliable and valid measures are therefore essential to increase understanding of such phenomena and inform appropriate treatment.

Method: A systematic review of the psychometric quality of existing VH measures was conducted using COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) methodology.

Results: Twenty articles were identified, with seventeen measures critically evaluated. Most measures did not provide a definition of VH or adequately explore psychometric properties, limiting their use in clinical and research settings. The VH-specific measure, the NEVHI demonstrated the highest quality evidence, with the MUSEQ, QPE, and PsyCH-Q also reporting adequate psychometrics, although further investigation into the reliability of such measures is needed.

Conclusions: Evaluation of the methodological strengths and weaknesses of the identified measures demonstrates the need for further validation studies, assessing multiple psychometric properties. Sufficient funding is required to ensure the robust development and content validity of new measures that incorporate the views of service-users to establish item relevance, comprehensiveness, and comprehensibility.

Word count: 178

Keywords: Visual Hallucinations; Measures; Reliability; Validity; Assessment

Introduction

Visual hallucinations (VH) are visual perceptions that occur, involuntarily, when awake, without the presence of an external stimulus (Aynsworth et al., 2017). Distinct from visual illusions or distortions which arise when stimuli in the immediate environment appear different to the expected perceptual experience (Waters et al., 2014), VHs are reported in a wide range of organic and psychiatric disorders, as well as non-clinical populations (Aynsworth et al., 2022; Schutte et al., 2020). Researched extensively in neurodegenerative conditions, up to 93% in individuals with Lewy Body Dementia (Aynsworth et al., 2017) and 75% of those with Parkinson's Disease (Urwyler et al., 2014) report experiencing VH. Such experiences are also common in visual disorders; typically occurring in up to 60% of patients with eye disease (Waters et al., 2014) and as many as 20% of individuals with severe glaucoma (Subhi et al., 2021). Although frequent VHs have traditionally been associated with schizophrenia spectrum disorders, with a recent prevalence rate of 33% in individuals with first episode psychosis (Allen et al., 2023) and 16-72% of people with schizoaffective disorder (Dudley et al., 2013), Schutte and colleagues (2020) reported the incidence and distress of such experiences as similar in borderline personality disorder (BDP) and post-traumatic stress disorder (PTSD), thus highlighting the shared phenomenology of visual hallucinations across conditions.

Complex VHs, such as people, animals, figures or faces, are most common in psychosis (van Ommen et al., 2019) as well as in non-clinical populations (Aynsworth et al., 2023). Unformed or simple VHs including flashes of light, shadows, unformed dots, and spirals are less typical (Oorschot et al., 2012). Frequently distressing, such experiences are associated with poorer treatment outcomes (Waters et al., 2014), increased hospital admissions and suicidality for patients with psychosis, (Chouinard et al., 2019; Yates et al., 2022) and more severe illness in neurodegenerative disorders (Clark et al., 2017).

Despite the detrimental impact of VHs, they have been largely overlooked in both the psychosis literature, and across conditions, and effective treatment options are lacking (Wilson et al., 2016). Existing measures examining the phenomenology of hallucinatory experiences have often not independently assessed individual modalities, (e.g. the Positive and Negative Syndrome Scale, PANSS; Kay et al., 1987) or have typically focused on the auditory hallucinatory domain, such as the Psychotic Symptom Rating Scales (PSYRATS; Haddock et al., 1999b). Accurate assessment of VHs is therefore vital to enhance understanding of such phenomena and develop appropriate treatment for those distressed by such experiences.

To the authors' knowledge, Aynsworth and colleagues conducted the only existing review of VH measures; 32 papers describing 19 measures were included (Aynsworth et al., 2017). The authors' development of a robust search strategy enabled their comprehensive review of three broad types of measure: those specific to VH, general hallucination measures, and psychosis instruments. However, this broad inclusion criteria meant that many of the reported instruments contained few VH specific items or no separate VH subscale, e.g., The Oxford-Liverpool Inventory of Feelings and Experiences, (Mason et al., 1995), Rating Scale for Psychotic Experience (Chouinard & Miller, 1999), Launay Slade Hallucination Scale (Launay & Slade, 1981), The Cardiff Anomalous Perceptions Scale (Bell et al., 2006). Specific, sensitive, reliable, and valid assessments of VH are required to aid understanding of such phenomena and inform treatment. Thus, it is pertinent that a review of VH measures focuses on VH specific instruments or those containing a separate VH subscale, given their ability to facilitate detailed exploration of these experiences. Furthermore, a bespoke evaluation grid was employed in the existing review to appraise the quality of the measures. As highlighted by the authors, this likely increased the subjectivity of their quality evaluation, thus reducing the scientific rigour of the review.

Considering these limitations, this review intended to update and expand upon Aynsworth et al.'s (2017) review by employing COSMIN (COnsensus-based Standards for the selection of health Measurement INstruments) methodology, the only consensus-based approach available to evaluate patient reported outcome measures (Mokkink et al., 2018; Mokkink et al., 2010a, 2010b), to assess the methodological and psychometric quality of existing VH specific measures in clinical and non-clinical populations. As well as providing clear definitions and guidance for assessing measurement properties, COSMIN methodology details criteria to assess the content validity of instruments, defined as *"the degree to which their content is an adequate reflection of the construct to be measured"* (p.7) (Terwee et al., 2018). Reported to be the most important measurement property for an instrument, it includes assessment of service-user consultation in establishing the relevance, comprehensiveness, and comprehensibility of questionnaire items. This is key to ensuring that a measure accurately captures the concepts of interest and is acceptable to the target population for which it is developed. Aynsworth et al. (2017) reported that only one measure, the North East Visual Hallucinations Interview (NEVHI; Mosimann et al., 2008), explored this measurement property when utilising their bespoke evaluation grid.

This review therefore aimed to systematically evaluate the content validity of the VH measures, by assessing the quality of each instrument's development study, and the content of the measure itself. Other primary properties evaluated within this domain were construct, structural, and criterion validity. Reliability, "*the extent to which scores for patients who have not changed are the same for repeated measurement under several conditions"* (p.11) (Mokkink et al., 2010b) was also assessed. Recommendations are included to aid the selection of effective measures for use in clinical practice and research.

Methods

Protocol and registration

The review was completed in accordance with Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, see Appendix 1.1 (Page et al., 2021). The protocol was registered on PROSPERO on 13/12/22, and the review's focus specified in an amendment on 16/07/23 (Registration number CRD42022370763, [available from: crd.york.ac.uk/prospero/display_record.php?ID=CRD42022370763:]).

Search Strategy

Scoping searches were initially conducted using the search terms outlined by Aynsworth et al. (2017). However, employing a broader range of search terms to describe outcome measures and aspects of validity and reliability, than those utilised by Aynsworth et al. (2017), resulted in the detection of more key papers. The amended search strategy was adapted for use with each database and reviewed by a University of Glasgow librarian (see Appendix 1.2 for an outline of the full search strategy). As the review aimed to update the existing one performed by Aynsworth et al (2017), published literature was retrieved from MEDLINE, EMBASE, CINAHL and PsycINFO databases from 7th November 2016 until the date of extraction, 13th June 2023. Given that an amended search strategy was employed in the current review, a search of the above databases, from inception until 7th November 2016, was also conducted to ascertain if the revised search terms identified any key papers not included in Aynsworth et al's. (2017) review. One relevant paper by Shine et al. (2015) was found and is included in the review.

Search results were exported to EndNote X9, and de-duplication was completed. Results were uploaded to Rayyan systematic reviewing software, where titles and abstracts were screened against the inclusion criteria.

Articles were initially screened for eligibility through title and abstract. The second reviewer (JM) independently rated 10% of the papers (N=183) for eligibility, which indicated near perfect reliability (99.5% agreement). For the few disagreements, the rationale for each rater's decision was discussed and a consensus reached. Studies included at this stage were then read in full and screened against eligibility criteria. Ten percent (N=3) of the full text papers were independently reviewed, indicating 100% agreement.

A manual search of the reference lists of included papers, and a forward-search of their citations, were completed, on 15th June 2023, to identify additional records. Two further eligible papers were found. Copies of eligible measures, and additional data, were sought by contacting the main authors of included papers between 19th-21st June 2023. The original development paper for the Questionnaire for Psychotic Experiences (QPE; Rossell et al., 2019) could not be accessed from the authors. A study examining the psychometric properties of the QPE and conducted by the measure's developers, was therefore graded as the development paper.

Eligibility Criteria

Inclusion criteria were studies: a) reporting novel instruments measuring VH, from 8th November 2016 onwards, or papers that explored additional psychometric properties of an existing VH measure. Instruments which measured hallucinations in other domains (e.g., auditory, olfactory, gustatory, tactile) were included if they contained a separate VH scale; b) for adults (16+) with organic or non-organic causes of VH (e.g., mental health or neurological disorder, cognitive difficulties, delirium or substance misuse) or non-clinical populations; c) published in the English language; d) full-text available in a peer-reviewed journal.

Exclusion criteria included: a) studies employing single items to assess VH; b) samples including individuals with physical health conditions (e.g., eye disease, such as Glaucoma or Charles Bonnet Syndrome). Papers that included such populations were considered if they also contained individuals with mental health diagnoses, neurological disorders, cognitive difficulties, delirium, substance misuse or non-clinical populations in the sample; c) case studies, unpublished dissertations, conference abstracts, or reviews.

Data Extraction

Characteristics of both the studies (e.g., sample, study setting, location) and included measures (e.g., mode of completion, number of VH items, structure, scoring, see appendix 1.3) were collated. Statistical and methodological information was extracted and evaluated to rate the quality of the measures using the manualised COSMIN approach, see Table 1.1 for COSMIN definitions of measurement properties (Mokkink et al., 2018; Mokkink et al., 2010a, 2010b; Prinsen et al., 2018; Terwee et al., 2018).

Table 1. 1: COSMIN definitions of measurement properties

Measurement Property	COSMIN Definition
Reliability	The degree to which the measurement is free from measurement error
Internal consistency	The degree of the interrelatedness among the items
Measurement error	The systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured
Test-Retest	The proportion of the total variance in the measurements which is because of "true" differences among patients
Validity	The degree to which an instrument measures the construct(s) it purports to measure
Content validity	The degree to which the content of an instrument is an adequate reflection of the construct to be measured
Hypothesis testing for Construct validity	The degree to which the scores of an instrument are consistent with hypotheses (for instance regarding internal relationships, relationships to scores of other instruments, or differences between relevant groups) based on the assumption that the instrument validly measures the construct to be measured
Structural validity	The degree to which the scores of an instrument are an adequate reflection of the dimensionality of the construct to be measured
Cross-cultural validity	The degree to which the performance of the items on a translated or culturally adapted instrument are an adequate reflection of the performance of the items of the original version of the instrument
Criterion validity	The degree to which the scores of an instrument are an adequate reflection of a "gold standard"
Responsiveness	The ability of an instrument to detect change over time in the construct to be measured

Quality assessment

Content Validity

COSMIN methodology was initially employed to evaluate the content validity of the measures (Mokkink et al., 2018b). Firstly, the quality of each measure's development was evaluated, by the completion of 35 items to assess its relevance and comprehensiveness. Quality ratings were based on predefined criteria in the COSMIN manual (Mokkink et al., 2018), and were rated as "very good", "adequate", "doubtful", "inadequate", or "not applicable" if the development criteria were not assessed. Item scores were pooled and an overall score for the measure's development study was determined using a "worse score counts" method (Terwee et al., 2012). Following this,

COSMIN guidelines stipulated that the quality of all content validity studies, those that evaluated the relevance, comprehensiveness, or comprehensibility of an existing instrument (Terwee et al., 2018), should be graded according to defined standards. As no content validity studies were found on any of the included VH measures, this step was not undertaken. The content validity of the VH subscale in each measure was then rated as sufficient (+), insufficient (-), or inconsistent (±) against 10 predetermined criteria (Terwee et al., 2007) (see appendix 1.6). In accordance with COSMIN guidelines, the main author also rated the content of each instrument against these criteria. These results were then qualitatively summarised to determine the overall content validity of each measure. Finally, a modified version of the Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) approach was used to determine overall evidence quality (Mokkink et al., 2018). “High” quality evidence was assumed, and the measures downgraded according to concerns regarding risk of bias (determined by the quality of each measure’s development study), inconsistency of results, or indirectness of evidence.

Methodological Quality of the included instruments

The COSMIN manual for evaluating the methodological quality and measurement properties of the included studies was utilised to grade the instruments. The risk of bias checklist (114 items separated into 10 boxes) aided methodological quality assessment through the evaluation of the strength of psychometric properties. Boxes 1 and 2 assessed instrument development and were completed in the content validity evaluation stage. Boxes 3 to 5 assessed aspects of internal structure (e.g., structural validity, internal consistency, cross-cultural validity) and 6 to 10 evaluated reliability (test-retest, inter-rater, measurement error), validity (criterion validity, hypothesis testing for construct validity) and responsiveness.

Similarly to the content validity assessment, methodological quality ratings were based on predefined criteria (Mokkink et al., 2018), and graded as “very good”, “adequate”, “doubtful”, “inadequate”, or “not applicable” if the property was not assessed. The “worse score counts” method (Terwee et al., 2012) was again applied to determine overall scores after all item scores were pooled. Twenty five percent of the included papers were independently rated by the second reviewer to test inter-rater reliability. There was 82.6% agreement between raters with discrepancies resolved through discussion.

Terwee (2007) and Prinsen (2016)’s criteria (see appendix 1.4) were employed to rate the quality of psychometric properties was sufficient (+), insufficient (-), or indeterminate (?). This uses standard statistical conventions, and incorporates some methodological requirements (e.g., for internal consistency to be rated as “sufficient”, evidence of structural validity is also required (Airey et al., 2023). Ratings were then quantitatively summarised for the measures that included further evaluation studies. Finally, the modified Grading of Recommendations Assessment, Development, and Evaluation (GRADE) approach was used to determine overall evidence quality (Mokkink et al., 2018). Downgrading of the evidence was again based on risk of bias, inconsistency of results, indirectness of evidence or imprecision (based on sample size).

Results

Overview of included studies

2713 records were identified through the database searches and a further 28 were found from the reference lists and citations of eligible articles. After duplicate removal and initial abstract/title screening, the full text of 55 records were reviewed, resulting in the inclusion of 9 studies (see Fig.1.1).

Eleven eligible studies from the previous review (Aynsworth et al., 2017) were identified; 8 were excluded due to their limited VH specific items (see appendix 1.5 for details of the excluded measures). A total of 20 studies were therefore included in the review and are summarised in Table 1.2.

The majority (n=18) of the studies were conducted in Western countries, primarily Europe (n=10), Australia (n=5) or America (n=3). A further 2 were undertaken in Asia; one in Japan and the other in Qatar. The sample sizes ranged from 15-1,300 (M=170.6, SD=277.37). One paper outlined a theoretical framework used to develop a novel hallucination measure, The Plausibility of Hallucinations Scale (PoH; Franceschi, 2011) but did not recruit a participant sample to assess its psychometric properties. By far the largest sample sizes were reported by Mitchell et al. (2017) in the creation of the Multi-Modality Unusual Sensory Experiences Questionnaire (MUSEQ; Mitchell et al., 2017); 513 for initial development and 659 for factor replication. 18 of the papers utilised clinical populations; 7 studies recruited individuals with a diagnosis of Parkinson's Disease (PD), 2 papers of participants with eye disease were included as they also recruited individuals with PD or cognitive impairment in the sample (Lai et al., 2016; Mosimann et al., 2008). A further 7 studies utilised samples with a psychiatric diagnosis, namely schizophrenia-spectrum disorder, Major Depression Disorder (MDD), Bipolar Affective Disorder (BD) or Borderline Personality Disorder (BPD). Two studies utilised non-clinical populations to investigate VHs. Kamp et al. (2023) recruited 310 widowed individuals who reported sensory experiences of their deceased partner. Although one study used a small clinical sample of individuals with a Schizophrenia or BD diagnosis for scale validation, students and community members participated in the larger development and replication study. Mean participant ages ranged from 20.92-79.9. Seventy percent (n=14) of the studies had a larger number of male to female participants. Two studies reported an equal gender proportion. Most studies were undertaken in outpatient

settings (n=13), 2 were in inpatient services and 4 combined inpatient and outpatient samples.

Figure 1.1: PRISMA flow diagram for identification of study

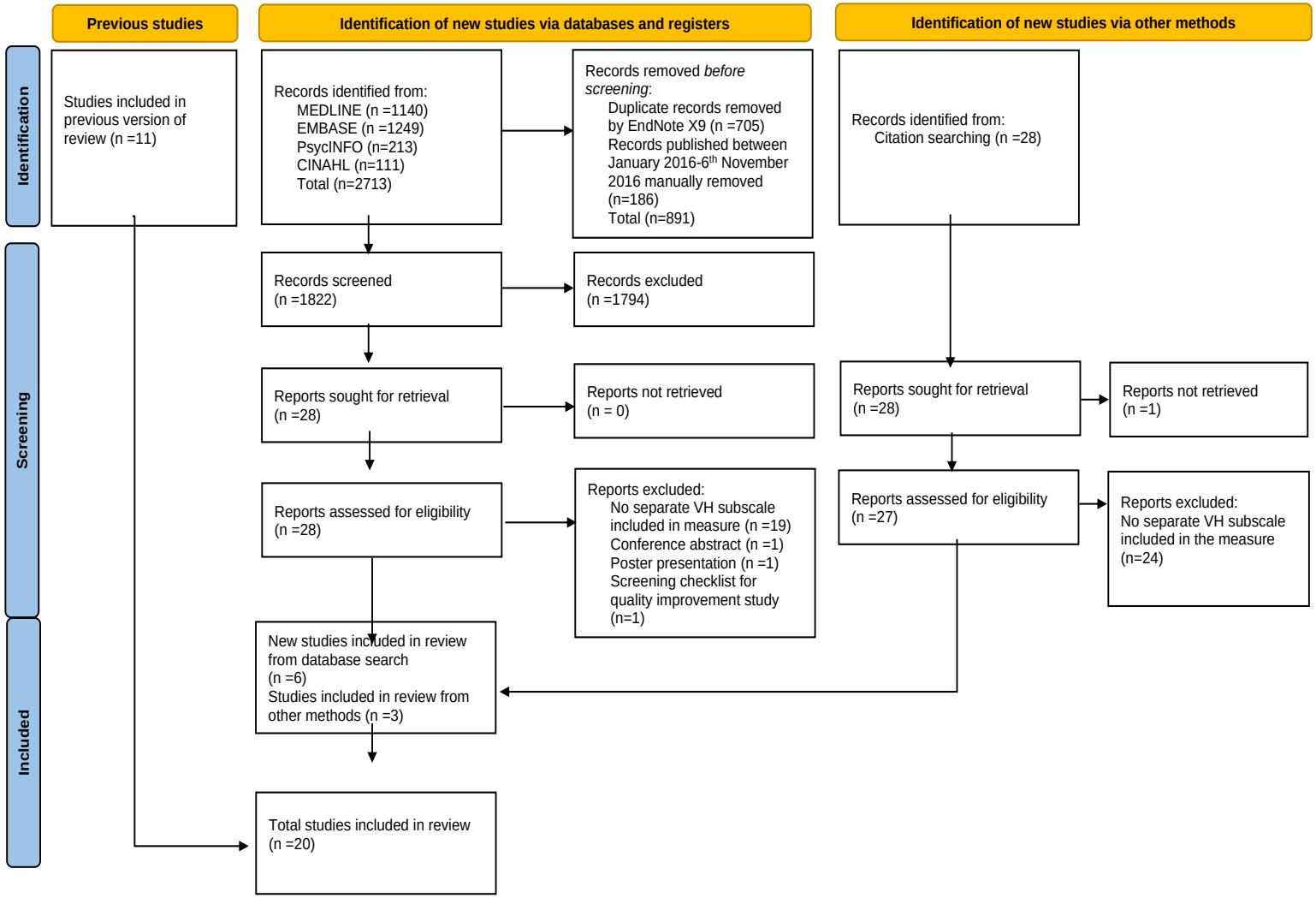

Table 1. 2: Overview of included articles

Study	Location	*N*	Population	Age M(SD)	Gender: Male/Female (%)	Setting	Visual Hallucination Measure
de Chazeron et al. (2015)	France	137	N:n=86; C:n=51	N:53.3 (19.5). C:38.5 (11.9)	N: 53.5: 46.5 C: 64.7: 35.3. Total sample: 62.8:37.2	Inpatient	Psycho-sensory Hallucinations Scale (PSAS)
Dudley et al. (2012)	England, UK	16	A	25 (4.31)	62.5:37.5	In and outpatient	Appraisals and Reactions to Visual Hallucinations Interview (ARVHI)
Dudley et al. (2013)	England, UK	176 (61 with VHs)	A	22 (4.5)	75:25	Outpatient	Psychotic Symptom Rating Scale-Visual Hallucinations (PSYRATS-VH)
Dudley et al. (2023)	England, UK	82	A	30.40 (10.34)	54.9:45.1	Outpatient	Multi-Modal Hallucinations Interview (MMHI)
Franceschi. (2011)	France	0	-	-	-	-	The Plausibility of Hallucinations Scale (PoH)
Gauntlett-Gilbert and Kuipers (2005)	England, UK	20	C, D, F, I, J, G	41.5 (12.4)	50:50	In and outpatient	Semi-Structured Interview about Visions in Psychiatric Patients (SSIVPP)

Study	**Location**	***N***	**Population**	**Age M(SD)**	**Gender: Male/Female (%)**	**Setting**	**Visual Hallucination Measure**
Holiday (2017)	America	177	N	68.9 (7.6)	50.8:49.2	Outpatient	North East Visual Hallucinations Interview (NEVHI)
Kamp et al. (2023)	Denmark	310	L	70.2 (8.28)	58.1:41.9	Outpatient	Sensory Experiences of Deceased Spouse (SED)
Lai et al. (2016)	England, UK	35	N:n=12, U:n=10, N&T:n=3, M:n=2, M&N: n=2, V&:n=2, M&T:n = 1, V&W:n = 1, W:n	71	68.6:31.4	Outpatient	Semi-Structured Interview of Complex Visual Hallucinations (SICVH)
Mitchell et al. (2017)	Australia	1,300: 513 (development) 32 (validation study) 659 (FA), 96 (TRR)	Development study: non-clinical. Validation study: B & F. FA & TRR: students	Development study: 27.75 (13.28) Validation study: 34.17 (13.09), FA: 20.92 (5.58), TRR: 19.82 (4.49)	Development study: 24.8:75.2 Validation study, clinical group: 37.5:62.7, FA: 33.5:66.5 TRR: 28:71: 1 (“other”)	Outpatient	The Multi-Modality Unusual Sensory Experiences Questionnaire (MUSEQ)
Muller et al. (2018)	Australia	163	N	69.7 (8.6)	58.2:41.8	Outpatient	Psychosis and Hallucinations Questionnaire (PsycH-Q)-Informant Version

Study	**Location**	***N***	**Population**	**Age M(SD)**	**Gender: Male/Female (%)**	**Setting**	**Visual Hallucination Measure**
Mosimann et al. (2008)	England, UK	114 older adults	S or U: n=80, Y: n=34	S or U: 79.9 (8.1), Y: 71.2 (8.7)	Clinical group: 32.5:67.5 Control group: 47.1: 52.9. Total sample: 36.8:63.2	Outpatient	North East Visual Hallucinations Interview (NEVHI)
Papapetropoulos et al. (2008)	America	70	N	Hallucinations: 63.9 (10.0), Without hallucinations: 64.3 (10.5)	Hallucinations 67.7:32.2, Without hallucinations: 64.1:35.9. Total sample: 65.7:43.3	Outpatient	The University of Miami Parkinson's Disease Hallucination Questionnaire (UM-PHQ)
Pappert et al. (1999)	America	126	N	57.4 (11.4)	58.7:41.3	Outpatient	Rush Hallucination Inventory (RHI)
Rossell et al. (2019)	Australia	173	C, D, F, H, Y	C: 40.3 (11.1), D: 43.4 (10.7), F: 32.1 (12.8), H: 30.2 (12.8), Y 28.6 (8.9)	C:44:56, D :51.6:48.4, F:51.6:48.4, H:41.2:58.8. Y:46.9 :53.1. Total sample: 43.4 : 56.6	In and outpatient	Questionnaire for Psychotic Experiences (QPE)
Shine et al. (2015)	Australia	197	N with hallucinations: n=86. N without hallucinations: n=111	With hallucinations: 70.5 (8.5), Without hallucinations: 68.6(8.4)	Total sample: 58.4:41.6, not specified by group	Outpatient	Psychosis and Hallucinations Questionnaire (PsycH-Q)

Study	**Location**	***N***	**Population**	**Age M(SD)**	**Gender: Male/Female (%)**	**Setting**	**Visual Hallucination Measure**
Van Ommen et al. (2019)	The Netherlands	45 (18 VHs, 10 VH-prone, 17 controls). 27 administered VH questionnaire	C:n=32, D:n=9, J:n=3, E:n=1	Total sample: 35.9 (9.5), VH and VH-prone group: 36.6 (9.5)	Total sample: 66.7:33.3. VH Sample: 55.6: 44.4	Outpatient	Visual Hallucination Questionnaire (VHQ)
Wada-Isoe et al. (2008)	Japan	41	N without dementia:n=3.N with dementia: n=10	N without dementia: 68.7 (10.1), N with dementia: 71.6 (7.7)	N without dementia: 25:874:2, N with dementia: 20:80. All participants: 24.4:75.6	Inpatient	Tottori University Hallucination Rating Scale (TUHARS)
Williams et al. (2008)	Australia	204	N:n=115; R:n= 22, Q:n=9,O: n=5, M: n=5, P:n= 19, Z:n= 15; Y:n=14	N:65, R:68.2, Q: 60.5, O: 74.8, M: 74.2, P: 66.6, Z:67.4, Y:62.6	N: 70:30, R: 59:41, Q: 33:67, O:20:80, M:80:20, P: 53:47, Z:53:47, Y:50:50	Outpatient	Queen Square Visual Hallucinations Interview (QSVHI)
Yehya et al. (2023)	Qatar	50	C:n=36, A:n=2, D:n=5, F:n=6, F&CL n=1	36.7 (12.6)	84:16	In and outpatient	Questionnaire for Psychotic Experiences (QPE)-Arab version

Population Code: A=Psychosis; B=Substance Induced Psychosis; C=Schizophrenia; D=Schizoaffective Disorder; E=Schizophreniform Disorder; F=Bipolar Affective Disorder; G=Borderline Personality Disorder; H=Major Depressive Disorder; I=Depression; J= Psychotic Disorder Not Otherwise Specified; K=Psychotic Depression; L= Widowed Individuals with sensory experiences of the deceased; M=Lewy Body Dementia; N=Parkinson's Disease; O=Vascular Parkinsonism; P= Unclassified Parkinsonism ; Q=Multiple System Atrophy; R= Progressive Supranuclear Palsy; S=Cognitive Impairment; T=Visual Impairment; U=Eye Disease; V=Brain Damage; W=Epilepsy; X=Unknown Cause; Y=Control; Z=Disease Control.

Psychometric properties: FA= Factor Analyses; TRR=Test-retest reliability; VH= Visual hallucinations.

Seventeen measures were reported in the studies and are summarised in appendix 1.3. Seven measures were specific to VH (NEVHI, PSYRATS-VH, ARVHI, QSVHI, SSIVPP, VHQ and SICVH). Seven instruments measured non-specific hallucinations but included a separate VH scale (PSAS, PoH, PsycH-Q, MMHI, TUHARS, UM-PHQ and RHI). Two measures were of unusual sensory experiences (MUSEQ and SED) but again contained VH scales, as did the QPE in its exploration of general psychotic-like symptoms. Most of the measures (*n*=12) broadly aimed to assess the presence of VH. Three measures had the specific purpose of investigating VH appraisals (ARVHI, SSIVPP, SICVH). Although reporting of the measures' aims was often limited, 2 detailed the intention of the instruments to be used in therapy to assess change in symptoms (PSYRATS-VH) or VH plausibility (PoH). Only 5 measures gave a definition of VH (QPE, NEVHI, SED, SICVH, VHQ). Nine (TUHARS, UM-PHQ, QSVHI, RHI, SSIVPP, MMHI, ARVHI, PSYRATS-VH, PsycH-Q) failed to provide a definition in relation to the phenomenon under exploration or specifically to VH. The lack of description of the construct to be measured was a major weakness of most of the instruments, and a threat to their content validity. This conceptual ambiguity was reflected by the content of the VH items in some of the measures. Questions relating to visual illusions, misperceptions, or felt presence were often included in VH subscales (e.g., in the PsycH-Q, MUSEQ, SICVH, and QSVHI) which undermines attempts to assess such phenomena and contributes to inconsistency in the literature (Aynsworth et al., 2017). The measures acquired a range of information on VH characteristics. General hallucination measures typically broadly assessed VH content and frequently, although the QPE also included single items investigating insight and the emotional and functional impact of such experiences. Unsurprisingly, the VH specific measures provided a more comprehensive assessment of the phenomenon. The ARVHI, SSIVPP and SICVH captured experiences in most detail, as seen in their larger number of VH items. These measures were the only to explore appraisals of VH. Likewise, both the ARVHI and SSIVPP assessed triggers and coping strategies, areas not covered by the other measures. Moreover, although the

association between trauma, in particular, repeated dissociative experiences, and visual hallucinations is well-documented (Quidé, 2023), few measures explored this connection and those that did, lacked detail on this area. The QPE and SSIVPP used a binary yes"/"no" option to ask respondents if their VH related to unpleasant past events. The SSIVPP also explored the presence of connected trauma symptoms, intrusive images, and memories. Regarding scoring, Likert rating scales were utilised in many measures to indicate differing levels of responses. However, 3 instruments (MMHI, PoH and QSVHI) only used binary "yes"/"no" options, thus limiting the information gleaned from responses.

Methodological Quality of measure development

Table 1.3 displays the risk of bias ratings for each measure's development study, see appendix 1.7 for content validity ratings and the rationale for the overall quality of evidence ratings. As mentioned, a serious failing in the design of 9 of the measures was their 'inadequate' description of the construct to be measured. Likewise, although most instruments gave a clear description of the rationale for their creation and intended target population, the proposed application of use was methodologically 'doubtful' or 'inadequate' for 10 of the measures. For many, little information was given on whether they were developed to evaluate interventions, for research purposes, or to explore prevalence rates. Notably, concept elicitation, which provides details on the relevance and comprehensiveness of instrument items was a methodological weakness of all the measures. Vague descriptions of the measures' development were often given, and it was not clear if an appropriate qualitative data collection method was employed in the studies to identify relevant items. Similarly, only 3 measures (SED, NEVHI and PycH-Q) performed a pilot test to assess comprehensibility and comprehensiveness. However, the methodological quality of these pilot tests was rated as 'doubtful' due to the lack of clarity regarding user involvement in assessing the comprehensibility of the measures. Fifteen (88.2%) of the measures' development studies were rated as 'inadequate.'

Fourteen of these received this rating due to their failure to perform a pilot or cognitive interview study. COSMIN guidance stipulates that a cognitive interview or other pilot study should be performed to evaluate the comprehensiveness and the comprehensibility of questionnaire items (Terwee et al., 2018). A cognitive interview is a qualitative survey development method used to evaluate how individuals process and respond to outcome measures (Meadows, 2021). It includes techniques such as verbal probes and “think aloud” protocols to support a researcher’s analysis of respondents’ comprehension of survey questions (Ryan et al., 2012). Moreover, the PsycH-Q’s lack of a clear definition of hallucinations resulted in its development study receiving an ‘inadequate’ methodological rating.

Table 1.3: Quality of Measure Development.

Measure	Measure design						Cognitive interview (CI) study[2]			TOTAL Measure DEVELOPMENT
	General design requirements					Concept elicitation	General design requirements	Comprehen-sibility	Total CI study	
	Clear construct	Clear origin of construct	Clear target population for which the Measure was developed	Clear context of use	Measure developed in sample represents the target population		CI study performed in sample representing the target population			
PSAS,(de Chazeron et al., 2015)	V	V	V	V	V	D	NR			I
ARVHI, (Dudley et al., 2012)	I	V	I	D	D	I	NR			I
PSYRATS-VH, (Dudley et al., 2013)	I	V	I	D	A	I	NR			I
MMHI, (Dudley et al., 2023)	I	V	V	I	D	D	NR			I

Note: Score: V= Very good; A = Adequate; D = Doubtful; I = Inadequate; NR= Not reported

Measure	Measure design						Cognitive interview (CI) study[2]			TOTAL Measure DEVELOPMENT
	General design requirements					Concept elicitation	General design requirements	Comprehen-sibility	Total CI study	
	Clear construct	Clear origin of construct	Clear target population for which the Measure was developed	Clear context of use	Measure developed in sample represents the target population		CI study performed in sample representing the target population			
PoH, Franceschi (2011)	V	V	V	I	V	I	NR			I
SSIVPP, (Gauntlett-Gilbert & Kuipers, 2005)	I	I	I	D	A	I	NR			I
SED, (Kamp et al., 2023)	V	V	V	D	V	D	V	D	D	D
SICVH (Lai et al., 2016)	V	V	I	D	A	I	NR			I
MUSEQ, Mitchell et al, (2017)	V	V	V	D	V	D	NR			I
NEVHI, (Mosimann et al., 2008)	V	V	V	V	V	D	D	D	D	D

Measure	Measure design						Cognitive interview (CI)			
	General design requirements					Concept elicitation	General design requirements	Comprehen-sibility	Total CI study	TOTAL Measure DEVELOPMENT
	Clear construct	Clear origin of construct	Clear target population for which the Measure was developed	Clear context of use	Measure developed in sample represents the target population		CI study performed in sample representing the target population			
UM-PHQ, (Papapetropoulos et al., 2008)	I	V	V	V	V	D	NR			I
RHI,(Papert et al., 1999)	I	V	V	D	V	D	NR			I
QPE, (Rossell et al., 2019)	V	V	V	V	V	D	NR			I
PsycH-Q, (Shine et al., 2015)	I	V	V	V	A	D	A	D	D	I
VHQ, (van Ommen et al., 2019)	V	V	I	D	D	I	NR			I
TUHARS, (Wada-Isoe et al., 2008)	I	V	V	V	V	I	NR			I
QSVHI, (Williams et al., 2008)	I	V	V	V	V	I	NR			I

Methodological Quality of the studies

Psychometric properties of only 6 measures ((MUSEQ, NEVHI, PSAS (de Chazeron et al., 2015), PsycH-Q (Shine et al., 2015), QPE, TUHRAS)) are reported in the development studies. Table 1.4 presents both the ratings for psychometric quality, and the overall quality of evidence (GRADE) rating, which is graded by methodological quality and consistency of psychometric evidence. Appendix 1.8 details the extracted study statistics which informed the ratings. Appendix 1.9 displays risk of bias ratings and the rationale for the overall quality of evidence ratings. The 7 studies that assessed internal consistency were methodologically 'very good.' Structural validity was measured as 'very good' or 'adequate' for 3 (75%) of the studies that assessed this psychometric property. However, one 'inadequate' methodology rating was reported due to the small sample size *(n*=41) in the PSAS development study. Likewise, criterion validity was measured consistently well in the 3 studies which evaluated this property. Construct validity assessed the quality for each comparator measure (convergent validity), or sample (discriminate validity) reported within the studies. 71.4% (*n*=5) of the methodological quality ratings were adequate or higher. Poor quality was seen for reliability which had 'doubtful' ratings in 66.7% (*n*=4) studies. This was mainly due to statistical reporting (e.g., weighted kappa not reported for ordinal scores) or an insufficient time interval between administrations of the measure (e.g., mean of 5.8 days between testing reported in the QPE; COSMIN guidelines specify that 2 weeks is considered appropriate for instrument evaluation (Mokkink et al., 2018; Prinsen et al., 2018; Terwee et al., 2018).

Table 1. 4: Psychometric quality and overall evidence quality ratings as defined by the GRADE approach.

Measure	*Content Validity*	*Structural Validity*	*Internal Consistency*	*Cross-Cultural Validity*	*Reliability*	*Measurement Error*	*Criterion Validity*	*Construct Validity/Hypothesis Testing*	*Responsiveness*
ARVHI, Dudley et al (2012)	+	NR	NR	NR	NR	NR	NR	NR	NR
Multi-Modal Hallucinations Interview, Dudley et al (2023)	+	NR	NR	NR	NR	NR	NR	NR	NR
MUSEQ, Mitchell et al (2017)	±	+	+	NR	-	NR	NR	+	NR
NEVHI, Mosimann et al, (2008)	+	+	+	NR	+	NR	NR	+	NR
PoH, Franceschi (2011)	+	NR	NR	NR	NR	NR	NR	NR	NR
PSAS, de Chazeron et al (2015)	+	?	-	NR	-	NR	-	NR	NR
PsycH-Q, Shine et al (2015)	±	NR	?	NR	+	NR	?	+	NR
PSYRATS-VH, Dudley et al (2013)	+	NR	NR	NR	NR	NR	NR	NR	NR
QPE, Rossell et al (2019)	+	+	+	NR	-	NR	-	?	NR
QSVHI, Williams et al (2008)	-	NR	NR	NR	NR	NR	NR	NR	NR
RHI, Pappert et al (1999)	±	NR	NR	NR	NR	NR	NR	NR	NR
SSIVPP, Gauntlett-Gilbert & Kuipers (2005)	+	NR	NR	NR	NR	NR	NR	NR	NR
Semi-structured interview of Complex Visual Hallucinations, Lai et al (2016)	+	NR	NR	NR	NR	NR	NR	NR	NR
SED, Kamp et al (2023)	+	NR	NR	NR	NR	NR	NR	NR	NR
TUHARS, Wada-Isoe et al (2008)	±	NR	?	NR	NR	NR	NR	+	NR
UM-PHQ, Papapetropoulos et al, (2008)	+	NR	NR	NR	NR	NR	NR	NR	NR
Visual Hallucination Questionnaire, Van Ommen et al (2019)	±	NR	NR	NR	NR	NR	NR	NR	NR

Quality of Evidence	Colour
High	
Moderate	
Low	
Very Low	

Appraisal of psychometric properties

Reliability

Internal Consistency: Subscale internal consistency was reported in 7 studies and ranged from 0.70 to 0.92. Alpha levels were all in the recommended range (Cronbach's α≥.70) apart from the PSAS's report of Kuder-Richardson alpha coefficients of .49 for the VH and .56 for the AH subscales. The measure's internal consistency as therefore rated as 'insufficient'. There was, however, 'sufficient' internal consistency within 3 measures (MUSEQ, NEVHI, QPE) alongside evidence of structural validity. Although adequate internal consistency was reported from the PsycH-Q's scores, structural validity statistics were not, thus the findings were rated as 'inconsistent.'

Inter-rater reliability: 4 studies reported inter-rater reliability. 'Sufficient' reliability (ICC or k≥.70) was reported in the clinician rated measures, the QPE and its Arabic version, as well as the NEVHI. The PSAS was rated as 'insufficient' due to reported Kappa statistics <.70.

Test-retest reliability: 4 studies provided intra-class correlation coefficients (ICC) or weighed Kappa statistics between .70-.92, indicating 'sufficient' test-retest reliability. The QPE, Arabic version, MUSEQ and PSAS were rated as insufficient' due to not meeting the specified statistical thresholds.

Validity

Content validity: As most of the measures' development studies were inadequate and there were no content validity studies available, the reviewers' ratings determined the overall content validity ratings (Prinsen et al., 2018). Content validity was rated as 'sufficient' in 64.7% of the studies. The QSVHI was graded as 'insufficient' due to the reviewers' concerns around the comprehensiveness and relevance of items to VH. Despite being a VH specific measure, only 3 of the 7 questions assessing hallucinatory experiences were focused on VH. Five measures received 'inconsistent' overall content

validity ratings, predominantly due to the vague reporting on the construct of interest and intended use of the measures.

Structural validity: 4 studies assessed structural validity. The QPE, NEVHI, and MUSEQ were ‘sufficient’, but the PSAS was rated as ‘indeterminate’ due the statistical reporting (e.g., comparative fit index not reported).

Criterion validity: Only 3 studies explored whether the measure’s score was an accurate reflection of a ‘gold standard’ instrument that assessed a similar construct. One study compared the QPE VH subscale to several psychotic symptom measures, e.g., Scale for the Assessment of Positive Symptoms (Andreasen, 1984), NEVHI (Mosimann et al., 2008) and the Positive and Negative Syndrome Scale (PANSS; Kay et al., 1987) and was rated ‘insufficient’ due to inter-scale correlation coefficients <.70. Likewise, comparing the PSAS to the Unified Parkinson’s Disease Rating Scale (UPDRS) and the PANSS in another study also found ‘insufficient’ validity from the available test scores. Another study compared the percentage of positive cases found by the clinician-rated Movement Disorder Society Unified Parkinson’s Disease Rating Scale (MDS-UPDRS; Goetz et al., 2008) with those identified by the self and informant PsyCH-Q (Muller et al., 2018). Statistical reporting (e.g., correlations or area under the receiver operating curve were not detailed in the paper) resulted in an ‘indeterminate’ rating.

Construct validity: Construct validity included convergent and discriminate validity. 5 studies assessed convergent validity. Most did not specify specific hypotheses so generic ones in relation to expected correlations with instruments measuring similar constructs were generated as per COSMIN guidelines (De Vet et al., 2011). Results were in accordance with these hypotheses for the NEVHI, MUSEQ, QPE and TUHARS which were rated as ‘sufficient.’

Discriminant validity was assessed in 4 studies. Results were in line with defined hypotheses in the MUSEQ and PsycH-Q development studies and so rated as ‘sufficient.’ The QPE was given an ‘indeterminate’ rating as discriminant validity was ‘sufficient’ in one study and ‘insufficient’ in the other.

Other Psychometric Properties

Notably, none of the studies explored other psychometric properties (e.g., cross-cultural validity, measurement error, responsiveness) detailed in the COSMIN guidelines. Although Yehra et al. (2003) tested psychometric properties of the Arabic version of the QPE, scores of two or more groups were not directly compared in one statistical model as is a COSMIN requirement for assessing cross-cultural validity (Prisen et al., 2018; Terwee et al., 2018; Mokkink et al., 2017). Therefore, this measurement property was not assessed in the Arabic QPE.

Quality of the evidence

Regarding overall evidence quality, 70% of evaluated measurement properties were rated as 'low' or 'very low.' All but 2 measures only had the development study report psychometric properties and the methodological quality, or risk of bias, often downgraded the evidence quality ratings. This was particularly evident for the studies that assessed test-retest reliability as well as for content validity, where the inadequate methodological quality of the measures' development studies led to a decrease in overall evidence gradings. Similarly, the sample size often resulted in a reduction in quality rating due to imprecision. Nine psychometric properties obtained high evidence quality ratings; these were predominately ratings of internal consistency**.**

Discussion

This review aimed to systematically evaluate the validity and reliability of existing visual hallucination measures. Twenty studies, detailing 17 measures were included; 7 of which were VH specific instruments and a further 10 assessed hallucinations, or unusual sensory experiences, more broadly but detailed a separate VH subscale. The development of current VH instruments was consistently poor and evidence for psychometric reliability and validity, only present for 6 of the 20 measures, was predominantly of low quality.

The current review enhances Aynsworth et al's (2017) in key areas. Although measures detailed in the existing review were included, a further 6 instruments (MUSEQ, MMHI, SICVH, SED, PsycH-Q, VHQ) were found and are reported. The use of a circumscribed eligibility criteria also ensured that only VH measures, or those with a separate VH subscale, were included, thus improving the specificity of the findings. This is pivotal for facilitating clinician and researcher decision making of the most appropriate measure for assessing VHs. Furthermore, this review employed COSMIN methodology, with clearly defined criteria to assess the overall quality of instruments. The use of such criteria to evaluate the content validity of the included measures was an evident strength of the review. It facilitated a robust evaluation of this important measurement property which included assessment of whether service-users were consulted in the development of the instruments. This is an evident advantage over the equivalent review which referred only to the views of *"experts in the field*" (p.178) in their assessment of content validity.

VH Specific Measures

As stated, the VH specific measures provided a more detailed and complex assessment of such phenomena. However, only the NEVHI explored psychometric properties; it's 'sufficient' ratings for structural and construct validity, internal consistency, and reliability resulted in it performing best of the measures reviewed. However, the overall evidence

quality of all measurement properties bar internal consistency was downgraded due to only one study of doubtful quality being available. Thus, further validation of this measure's psychometric properties is warranted.

General hallucination and psychotic symptom measures

Eight measures assessed general hallucinatory experiences but included a separate VH subscale; 4 provided psychometric evaluation of the novel instrument. The QPE showed most promising results out of these measures, receiving 'sufficient' structural validity and internal consistency as well as multiple high quality evidence ratings. This measure also concisely investigated a broad spectrum of hallucinatory characteristics, highlighting its utility in clinical settings. Further research into the psychometric properties of the QPE in larger samples is, however, recommended. Furthermore, out of the 4 measures included that were designed to assess hallucinations in PD populations (TUHARS, RHI, QSVHI, PsyCH-Q), the PsyCH-Q was the only one to report psychometrics. Overall quality ratings were 'high' or 'moderate' across all bar one psychometric property reported. Reliability and construct validity were rated as 'sufficient'. The lack of evidence for the PsyCH-Q's structural validity impacted its internal consistency rating, however this was noted to be >.70 in the development study reviewed. The measure's criterion validity, assessed in Muller et al's (2018) validation study, also achieved an 'indeterminate' psychometric rating due to the small sample size and unclear statistical reporting.

Unusual Sensory Experiences (USE) measures

The SED and MUSEQ were developed to explore unusual sensory experiences in clinical and non-clinical populations. The MUSEQ reported encouraging results with overall 'high' quality ratings and 'sufficient' psychometric properties for structural validity, internal consistency, and construct validity. Test-retest reliability was, however, 'insufficient' due to levels <.70. Again, the presence of only one small *n* study assessing

psychometric properties impacted the reliability quality rating, thus demonstrating the need for further research into this relatively newly developed measure. Moreover, along with the NEVHI, the very recently published SED detailed the most thorough approach to item construction and testing, as is reflected in the higher overall quality rating for content validity. Designed to assess USE of the deceased, the SED adds to the felt presence literature, and could prove promising with research into its psychometric properties.

Limitations

Although the COSMIN methodology provides clear and comprehensive criterion to determine the methodological and psychometric quality of instruments, it has several noteworthy limitations which affected the reviewers' appraisal of the included papers. The conservative nature of the COSMIN approach has been critiqued by several authors (Airey et al., 2023; Justo-Núñez et al., 2022; Smith et al., 2021). The use of the worst score counts principle (Terwee et al., 2012), in particular, downgraded both measurement design and methodological quality ratings, resulting in a more cautious appraisal of the studies. For example, as highlighted by Smith et al. (2019), the criteria for hypothesis testing appeared excessively stringent. Measures that did not specify priori hypotheses were downgraded even when the expected direction of the predicted effect was obvious and therefore unnecessary to state explicitly.

Moreover, like Airey et al. (2023), the reviewers of the current study frequently found that many of the included articles reported a lack of detail, especially in relation to the measures' development. Some studies may, therefore, have met COSMIN requirements but possible word count constraints could have demoted quality ratings (Airey et al., 2023). This is evidenced by the methodological quality rating of the SED's development study, which achieved a higher rating due to the reviewer's reference to supplementary information detailing further information on the design phases. Equivalent information was sought, but not obtained, from all other authors.

The absence of any content validity studies, and almost uniform inadequacy of the included measures' development, resulted in overall content validity ratings being determined by the reviewers as is in meeting with COSMIN recommendations (Terwee et al., 2018). This however increased the role for subjective evaluation. Likewise, the risk of publication bias was not accounted for when grading the quality of evidence as was deemed difficult to assess this due to the absence of registries for studies on measurement properties (Mokkink et al., 2018). An overestimation of the psychometric properties of the included measures may therefore have occurred in this review, given that studies with positive results are more likely to be published (Justo-Núñez et al., 2022).

Lastly, the exclusion of grey literature and non-English language studies potentially excluded other important measures or additional information on the psychometric properties of included measures, e.g., cross-cultural validity. Although Mokkink et al. (2018) recommend that language restrictions are not applied as part of the search strategy, appraisal of non-English language studies was beyond the scope of this review.

Implications for future research

Aynsworth et al. (2017) concluded that the lack of VH specific measures emphasised the need for further instrument development to increase understanding of hallucinatory phenomena. However, the absence of content validity studies and overriding inadequacy of the measures' development, suggests that before novel instruments are developed, further qualitative research (e.g., cognitive interviews) into existing measures must be undertaken, to identify whether items are relevant, comprehensive, and comprehensible to individuals experiencing VHs. This is especially pertinent given the lack of integration of items assessing the relation of hallucinatory experiences to traumatic experiences. However, qualitative research into scale development which includes psychology, psychiatry, neurology, psychometrician, and researcher

collaboration, as well as stakeholder consultation, requires adequate funding. Interestingly, only 8 (40%) of the included studies reported obtaining financial support to develop their measure.

Evidently, further research is required into all measures reported in the review. Internal consistency data was more commonly evidenced but adequate reliability information was sparse across the included papers. Studies including properties not evaluated (e.g., responsiveness and cross-cultural validity) are also urgently required. The lack of cultural exploration of measures is concerning, given that psychotic disorders are more common in ethnic minority groups (Jongsma et al., 2021) and VH has been reported to be more prevalent in non-Western populations (Ghanem et al., 2023). Culture has been found to influence the formation, maintenance, and explanation of hallucinations (Ghanem et al., 2023; Taitimu et al., 2018). The involvement of diverse stakeholder groups to establish the cross-cultural validity of measures is therefore imperative.

Conclusions

The current review critically evaluated the reliability and validity of current VH measures using COSMIN methodology. Expanding upon Aynsworth et al's (2017) findings, this review provides an updated compilation of existing instruments to support clinicians and researchers' assessment of such phenomena in different populations. Although the review identified 17 measures which described a range of VH phenomenology, most did not assess psychometric properties. It was, therefore, not possible to comment on their suitability for use in clinical practice or research. Furthermore, as highlighted by Aynsworth et al (2017), the absence of a clear definition of VH in most measures hinders the assessment of such phenomena and results in inconsistency of findings. The NEVHI demonstrated the most proficient psychometrics for the detailed assessment of VH specific phenomenology. The MUSEQ also reported good validity for the transdiagnostic investigation of VH, and so could be considered a useful screening tool in clinical settings. However further research into its reliability is warranted. Similarly, the QPE

evidences encouraging psychometric properties and can facilitate a detailed assessment of hallucinatory experiences, including their impact and relation to adverse events. Considering instruments that assess VH in PD populations, the PsycH-Q was the only measure to report adequate psychometrics. Clearly more comprehensive evaluation of existing measures is required, in large scale studies which investigate multiple psychometric variables. However, consideration must be given to the robust design of instrument items; this includes ongoing consultation with the target population and sufficient funding to examine relevance, comprehensiveness, and comprehensibility.

References

Airey, N. D., Taylor, C. D. J., Vikram, A., & Berry, K. (2023). Trauma measures for use with psychosis populations: A systematic review of psychometric properties using COSMIN. *Psychiatry Res*, *323*, 115163. https://doi.org/10.1016/j.psychres.2023.115163

Allen, S., Goodall, T., Jones, C., James, R., & Surtees, A. (2023). What Is the Prevalence of Visual Hallucinations in a First-Episode Psychosis Population? A Systematic Review and Meta-analysis of the Literature. *Schizophrenia Bulletin Open*, *4*(1). https://doi.org/10.1093/schizbullopen/sgad002

Andreasen, N. C. (1984). Scale for the assessment of positive symptoms. *Psychiatrie & Psychobiologie*.

Aynsworth, C., Collerton, D., & Dudley, R. (2017). Measures of visual hallucinations: Review and recommendations [Empirical Study; Literature Review; Qualitative Study]. *Clinical Psychology Review*, *57*, 164-182. https://doi.org/https://dx.doi.org/10.1016/j.cpr.2017.05.001

Aynsworth, C., Rolinson, J., Pervez, M., Collerton, D., & Dudley, R. (2022). What is the frequency and nature of visual hallucinations in non-clinical participants? *Psychology and psychotherapy.*, *11*. https://doi.org/https://dx.doi.org/10.1111/papt.12440

Bell, V., Halligan, P. W., & Ellis, H. D. (2006). The Cardiff Anomalous Perceptions Scale (CAPS): a new validated measure of anomalous perceptual experience. *Schizophrenia Bulletin*, *32*(2), 366-377.

Chouinard, G., & Miller, R. (1999). A rating scale for psychotic symptoms (RSPS): Part I: theoretical principles and subscale 1: perception symptoms (illusions and hallucinations). *Schizophrenia Research*, *38*(2-3), 101-122.

Chouinard, V.-A., Shinn, A. K., Valeri, L., Chouinard, P. A., Gardner, M. E., Asan, A., Cohen, B. M., & Ongur, D. (2019). Visual hallucinations associated with multimodal hallucinations, suicide attempts and morbidity of illness in psychotic disorders [Empirical Study; Quantitative Study]. *Schizophrenia Research*, *208*, 196-201. https://doi.org/https://dx.doi.org/10.1016/j.schres.2019.02.022

Clark, M. L., Waters, F., Vatskalis, T. M., & Jablensky, A. (2017). On the interconnectedness and prognostic value of visual and auditory hallucinations in first-episode psychosis. *European Psychiatry*, *41*, 122-128. https://doi.org/https://doi.org/10.1016/j.eurpsy.2016.10.011

de Chazeron, I., Pereira, B., Chereau-Boudet, I., Brousse, G., Misdrahi, D., Fenelon, G., Tronche, A. M., Schwan, R., Lancon, C., Marques, A., Debilly, B., Durif, F., & Llorca, P. M. (2015). Validation of a Psycho-Sensory hAllucinations Scale (PSAS) in schizophrenia and Parkinson's disease. *Schizophrenia Research*, *161*(2), 269-276. https://doi.org/doi:https://dx.doi.org/10.1016/j.schres.2014.11.010

Dudley, R., Collerton, D., Nicholson, M., & Mosimann, U. (2013). Clinical characteristics of disclosed visual hallucinations in users of an Early Intervention in Psychosis Service [Empirical Study; Interview; Quantitative Study]. *Psychosis: Psychological, Social and Integrative Approaches*, *5*(2), 127-133. https://doi.org/https://dx.doi.org/10.1080/17522439.2012.699543

Dudley, R., Watson, F., O'Grady, L., Aynsworth, C., Dodgson, G., Common, S., Day, B. A., & Fernyhough, C. (2023). Prevalence and nature of multi-sensory and multi-modal hallucinations in people with first episode psychosis. *Psychiatry Research*, *319*, 114988. https://doi.org/doi:https://dx.doi.org/10.1016/j.psychres.2022.114988

Dudley, R., Wood, M., Spencer, H., Brabban, A., Mosimann, U. P., & Collerton, D. (2012). Identifying specific interpretations and use of safety behaviours in people with distressing visual hallucinations: an exploratory study. *Behav Cogn Psychother*, *40*(3), 367-375. https://doi.org/10.1017/s1352465811000750

Franceschi, P. (2011). On the plausibility of psychotic hallucinations. *Activitas Nervosa Superior*, *53*, 35-43. https://doi.org/doi:http://dx.doi.org/10.1007/BF03379932

Gauntlett-Gilbert, J., & Kuipers, E. (2005). Visual hallucinations in psychiatric conditions: appraisals and their relationship to distress. *British Journal of Clinical Psychology*, *44*(1), 77-87. https://doi.org/doi:10.1348/014466504x19451

Ghanem, M., Evangeli-Dawson, C., & Georgiades, A. (2023). The role of culture on the phenomenology of hallucinations and delusions, explanatory models, and help-seeking attitudes: A narrative review. *Early Intervention in Psychiatry*, *17*(9), 843-863. https://doi.org/https://doi.org/10.1111/eip.13449

Goetz, C. G., Tilley, B. C., Shaftman, S. R., Stebbins, G. T., Fahn, S., Martinez-Martin, P., Poewe, W., Sampaio, C., Stern, M. B., Dodel, R., Dubois, B., Holloway, R., Jankovic, J., Kulisevsky, J., Lang, A. E., Lees, A., Leurgans, S., LeWitt, P. A., Nyenhuis, D., . . . LaPelle, N. (2008). Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Mov Disord*, *23*(15), 2129-2170. https://doi.org/10.1002/mds.22340

Haddock, G., McCarron, J., Tarrier, N., & Faragher, E. B. (1999). Scales to measure dimensions of hallucinations and delusions: the psychotic symptom rating scales (PSYRATS). *Psychological Medicine*, *29*(4), 879-889. https://doi.org/10.1017/S0033291799008661

Jongsma, H. E., Karlsen, S., Kirkbride, J. B., & Jones, P. B. (2021). Understanding the excess psychosis risk in ethnic minorities: the impact of structure and identity. *Soc Psychiatry Psychiatr Epidemiol*, *56*(11), 1913-1921. https://doi.org/10.1007/s00127-021-02042-8

Justo-Núñez, M., Morris, L., & Berry, K. (2022). Self-report measures of secure attachment in adulthood: A systematic review. *Clin Psychol Psychother*, *29*(6), 1812-1842. https://doi.org/10.1002/cpp.2756

Kamp, K. S., Steffen, E. M., Moskowitz, A., & Spindler, H. (2023). Prevalence and Phenomenology of Sensory Experiences of a Deceased Spouse: A Survey of Bereaved Older Adults. *Omega - Journal of Death & Dying*, *87*(1), 103-125. https://doi.org/https://dx.doi.org/10.1177/00302228211016224

Kay, Fiszbein, A., & Opler, L. A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin*, *13*(2), 261-276.

Lai, S., Bruce, V., & Collerton, D. (2016). Visual Hallucinations in Older People: Appraisals but not Content or Phenomenology Influence Distress. *Behavioural and Cognitive Psychotherapy*, *44*, https://ovidsp.ovid.com/ovidweb.cgi?T=JS&CSC=Y&NEWS=N&PAGE=fulltext&D=emed17&AN=614316442

Launay, G., & Slade, P. (1981). The measurement of hallucinatory predisposition in male and female prisoners. *Personality and Individual Differences*, *2*(3), 221-234. https://doi.org/https://doi.org/10.1016/0191-8869(81)90027-1

Mason, O., Claridge, G., & Jackson, M. (1995). New scales for the assessment of schizotypy. *Personality and Individual Differences*, *18*(1), 7-13.

Meadows, K.(2021). Cognitive Interviewing Methodologies. *Clinical Nursing Research*. *30(4)*:375-379. doi:10.1177/10547738211014099

Mitchell, C. A. A., Maybery, M. T., Russell-Smith, S. N., Collerton, D., Gignac, G. E., & Waters, F. (2017). The Structure and Measurement of Unusual Sensory Experiences in Different Modalities: The Multi-Modality Unusual Sensory Experiences Questionnaire (MUSEQ) [Original Research]. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.01363

Mokkink, L. B., De Vet, H. C., Prinsen, C. A., Patrick, D. L., Alonso, J., Bouter, L. M., & Terwee, C. B. (2018). COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Quality of Life Research*, *27*, 1171-1179.

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & De Vet, H. C. (2010a). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of Life Research*, *19*, 539-549.

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. (2010b). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of clinical epidemiology*, *63*(7), 737-745.

Mosimann, U. P., Collerton, D., Dudley, R., Meyer, T. D., Graham, G., Dean, J. L., Bearn, D., Killen, A., Dickinson, L., Clarke, M. P., McKeith, I. G., Mosimann, U. P., Collerton, D., Dudley, R., Meyer, T. D., Graham, G., Dean, J. L., Bearn, D., Killen, A., & Dickinson, L. (2008). A semi-structured interview to assess visual hallucinations in older people. *International Journal of Geriatric Psychiatry*, *23*(7), 712-718. https://doi.org/doi:10.1002/gps.1965

Muller, A. J., Mills, J. M. Z., O'Callaghan, C., Naismith, S. L., Clouston, P. D., Lewis, S. J. G., & Shine, J. M. (2018). Informant- and Self-Appraisals on the Psychosis and Hallucinations Questionnaire (PsycH-Q) Enhances Detection of Visual Hallucinations in Parkinson's Disease. *Movement Disorders Clinical Practice*, *5*(6), 607-613. https://doi.org/https://dx.doi.org/10.1002/mdc3.12683

Oorschot, M., Lataster, T., Thewissen, V., Bentall, R., Delespaul, P., & Myin-Germeys, I. (2012). Temporal dynamics of visual and auditory hallucinations in psychosis. *Schizophrenia Research*, *140(1-3)*, 77-82. https://doi.org/https://dx.doi.org/10.1016/j.schres.2012.06.010

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., . . . Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, *372*, n71. https://doi.org/10.1136/bmj.n71

Papapetropoulos, S., Katzen, H., Schrag, A., Singer, C., Scanlon, B. K., Nation, D., Guevara, A., & Levin, B. (2008). A questionnaire-based (UM-PDHQ) study of

hallucinations in Parkinson's disease. *BMC Neurology*, *8*. https://doi.org/doi:https://dx.doi.org/10.1186/1471-2377-8-21

Pappert, E. J., Goetz, C. G., Niederman, F. G., Raman, R., & Leurgans, S. (1999). Hallucinations, sleep fragmentation, and altered dream phenomena in Parkinson's disease. *Mov Disord*, *14*(1), 117-121. https://doi.org/10.1002/1531-8257(199901)14:1

Prinsen, Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., de Vet, H. C. W., & Terwee, C. B. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research*, *27*(5), 1147-1157. https://doi.org/10.1007/s11136-018-1798-3

Quidé, Y. (2023). Dissociation, trauma and the experience of visual hallucinations in post-traumatic stress disorder and schizophrenia. *BJPsych Open*, *9*(1), e23. https://doi.org/10.1192/bjo.2023.3

Rossell, S. L., Schutte, M. J. L., Toh, W. L., Thomas, N., Strauss, C., Linszen, M. M. J., van Dellen, E., Heringa, S. M., Teunisse, R., Slotema, C. W., & Sommer, I. E. C. (2019). The Questionnaire for Psychotic Experiences: An Examination of the Validity and Reliability. *Schizophrenia Bulletin*, *45*(45), S78-S87. https://doi.org/doi:https://dx.doi.org/10.1093/schbul/sby148

Ryan, K., Gannon-Slater, N., & Culbertson, M. J. (2012). Improving Survey Methods With Cognitive Interviews in Small-and Medium-Scale Evaluations. *American Journal of Evaluation*, *33*(3), 414-430. https://doi.org/10.1177/1098214012441499

Schutte, M. J., Linszen, M. M., Marschall, T. M., ffytche, D. H., Koops, S., van Dellen, E., Heringa, S. M., Slooter, A. J., Teunisse, R., van den Heuvel, O. A., Lemstra, A. W., Foncke, E. M., Slotema, C. W., de Jong, J., Rossell, S. L., & Sommer, I. E. (2020). Hallucinations and other psychotic experiences across diagnoses: A comparison of phenomenological features [Empirical Study; Quantitative Study]. *Psychiatry Research Vol 292 2020, ArtID 113314*, *292*. https://doi.org/https://dx.doi.org/10.1016/j.psychres.2020.113314

Shine, J. M., Mills, J. M. Z., Qiu, J., O'Callaghan, C., Terpening, Z., Halliday, G. M., Naismith, S. L., & Lewis, S. J. G. (2015). Validation of the Psychosis and Hallucinations Questionnaire in Non-demented Patients with Parkinson's Disease. *Movement Disorders Clinical Practice*, *2(2)*, 175-181. https://doi.org/https://dx.doi.org/10.1002/mdc3.12139

Smith, E. L., Garety, P. A., Harding, H., & Hardy, A. (2021). Are there reliable and valid measures of anxiety for people with psychosis? A systematic review of psychometric properties. *Psychol Psychother*, *94*(1), 173-198. https://doi.org/10.1111/papt.12265

Subhi, Y., Schmidt, D. C., Bach-Holm, D., Kolko, M., & Singh, A. (2021). Prevalence of Charles Bonnet syndrome in patients with glaucoma: a systematic review with meta-analyses [Review]. *Acta Ophthalmologica*, *99(2)*, 128-133. https://doi.org/https://dx.doi.org/10.1111/aos.14567

Taitimu, M., Read, J., & McIntosh, T. (2018). Ngā Whakāwhitinga (standing at the crossroads): How Māori understand what Western psychiatry calls “schizophrenia”. *Transcultural Psychiatry*, *55*(2), 153-177. https://doi.org/10.1177/1363461518757800

Terwee, C. B., Bot, S. D., de Boer, M. R., van der Windt, D. A., Knol, D. L., Dekker, J., Bouter, L. M., & de Vet, H. C. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of clinical epidemiology*, *60*(1), 34-42.

Terwee, C. B., Mokkink, L. B., Knol, D. L., Ostelo, R. W., Bouter, L. M., & de Vet, H. C. (2012). Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Quality of Life Research*, *21*, 651-657.

Terwee, C. B., Prinsen, C. A. C., Chiarotto, A., Westerman, M. J., Patrick, D. L., Alonso, J., Bouter, L. M., de Vet, H. C. W., & Mokkink, L. B. (2018). COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Quality of Life Research*, *27*(5), 1159-1170. https://doi.org/10.1007/s11136-018-1829-0

Urwyler, P., Nef, T., Killen, A., Collerton, D., Thomas, A., Burn, D., McKeith, I., & Mosimann, U. P. (2014). Visual complaints and visual hallucinations in Parkinson's disease. *Parkinsonism and Related Disorders*, *20(3)*, 318-322. https://doi.org/https://dx.doi.org/10.1016/j.parkreldis.2013.12.009

van Ommen, M., van Laar, T., Cornelissen, F., & Bruggeman, R. (2019). Visual hallucinations in psychosis. *Psychiatry Research Vol 280 2019, ArtID 112517*, *280*. https://doi.org/doi:https://dx.doi.org/10.1016/j.psychres.2019.112517

Wada-Isoe, K., Ohta, K., Imamura, K., Kitayama, M., Nomura, T., Yasui, K., Nakaso, K., & Nakashima, K. (2008). Assessment of hallucinations in Parkinson's disease using a novel scale. *Acta Neurologica Scandinavica*, *117(1)*, 35-40. https://doi.org/https://dx.doi.org/10.1111/j.1600-0404.2007.00907.x

Waters, F., Collerton, D., Ffytche, D. H., Jardri, R., Pins, D., Dudley, R., Blom, J. D., Mosimann, U. P., Eperjesi, F., Ford, S., & Laroi, F. (2014). Visual hallucinations in the psychosis spectrum and comparative information from neurodegenerative disorders and eye disease. *Schizophrenia Bulletin*, *40(SUPPL. 4)*, S233-S245. https://doi.org/https://dx.doi.org/10.1093/schbul/sbu036

Williams, D. R., Warren, J. D., & Lees, A. J. (2008). Using the presence of visual hallucinations to differentiate Parkinson's disease from atypical parkinsonism. *Journal of Neurology, Neurosurgery and Psychiatry*, *79(6)*, 652-655. https://doi.org/https://dx.doi.org/10.1136/jnnp.2007.124677

Wilson, R., Collerton, D., Freeston, M., Christodoulides, T., & Dudley, R. (2016). Is Seeing Believing? The Process of Change During Cognitive-behavioural Therapy for Distressing Visual Hallucinations [Case Reports]. *Clinical Psychology & Psychotherapy*, *23*(4), 285-297. https://doi.org/https://dx.doi.org/10.1002/cpp.1962

Yates, K., Lång, U., Peters, E. M., Wigman, J. T. W., Boyda, D., McNicholas, F., Cannon, M., Alderson-Day, B., Bloomfield, M., Ramsay, H., & Kelleher, I. (2022). Hallucinations as a risk marker for suicidal behaviour in individuals with a history of sexual assault: a general population study with instant replication. *Psychol Med*, 1-7. https://doi.org/10.1017/s0033291722001532

Chapter 2

Exploring the phenomenology of hallucinatory experiences and their relation to trauma re-experiencing symptoms: The Questionnaire of Hallucinatory Experiences (QHE)

Prepared in accordance with Frontiers in Psychology:
https://www.frontiersin.org/guidelines/author-guidelines

Authors:
Julia Caird[1], Professor Andrew Gumley[2], Dr. Kirsten Atherton[3], Dr. Andy Siddaway[4]

1. Mental Health and Wellbeing, School of Health and Wellbeing, University of Glasgow; and NHS Greater Glasgow and Clyde
2. Mental Health and Wellbeing, School of Health and Wellbeing, University of Glasgow; and NHS Greater Glasgow and Clyde
3. NHS Greater Glasgow and Clyde, UK
4. NHS Lanarkshire, UK

Corresponding address:
Clarice Pears Building
University of Glasgow
90 Byres Road
Glasgow
G12 8TB
j.caird.1@research.gla.ac.uk

Word count: 9,900

Plain Language Summary

Title

Developing a questionnaire of hallucinations across different sensory modalities and exploring overlapping hallucinatory experiences in trauma and psychosis.

Background

The link between trauma and hallucinations, when an individual sees, hears, smells, tastes, feels, or senses phenomena not experienced by others, is well established (Varese et al., 2012). Limited studies have investigated the link between trauma and psychotic symptoms, such as hallucinations. Hallucinations, once viewed as symptomatic of psychosis, have been reported across different conditions and especially with people with Post-Traumatic Stress Disorder (PTSD; Schutte et al., 2020). Understanding these hallucinatory experiences is crucial to informing treatment interventions. At present, no measures exist that assess the range of hallucinatory experiences found in trauma and psychosis.

Aims

The study aimed to develop a questionnaire of hallucinatory phenomena which captures the range of experiences reflected in trauma and psychosis. Three phases were employed:

- Phase 1: The research team generated an exhaustive list of hallucinatory experiences, used to create questionnaire items.
- Phase 2: Key researchers, clinicians, and individuals with lived experience of hallucinations reviewed and refined the items.
- Phase 3: Individuals with lived experience completed the questionnaire in a pilot to establish next steps for its development. Responses were analysed according to novel “go/revise/stop” criteria, created by the research team to assess the psychometric properties of the questionnaire.

Methods

Individuals with lived experience of hallucinations were recruited online from social media sites and special interest groups, to participate in phase 3 of the study. Individuals were invited to complete the questionnaire after gaining informed consent and answering screening questions to establish their eligibility.

Main Findings

Participants reported frequently experiencing a range of hallucinations which were distressing and impacted their functioning. Individuals reported that their hallucinations were linked to traumatic experiences but scores on the trauma scales were weakly correlated to the hallucinatory scales. Internal consistency was moderate, suggesting that the items in the questionnaire may not consistently measure the same thing. The results were in line with the “revise” criteria, indicating that further revision is required to the QHE.

Conclusions

Further development of the QHE is required and should include psychological theory of models used to treat psychosis and trauma. Including people with lived experience of hallucinations and trauma in the design, and testing, of questionnaire items is essential. We hope the QHE will support a shared understanding of hallucinations and their link to trauma and victimisation.

References

Schutte, M. J. L., Linszen, M. M. J., Marschall, T. M., ffytche, D. H., Koops, S., van Dellen, E., Heringa, S. M., Slooter, A. J. C., Teunisse, R., van den Heuvel, O. A., Lemstra, A. W., Foncke, E. M. J., Slotema, C. W., de Jong, J., Rossell, S. L., & Sommer, I. E. C. (2020). Hallucinations and other psychotic experiences across diagnoses: A comparison of phenomenological features. *Psychiatry Research*, *292*.

Varese, F., Smeets, F., Drukker, M., Lieverse, R., Lataster, T., Viechtbauer, W., Read, J., van Os, J., & Bentall, R. P. (2012). Childhood adversities increase the risk of psychosis: A meta-analysis of patient-control, prospective-and cross-sectional cohort studies. *Schizophrenia Bulletin*, *38*(4), 661–671.

Abstract

Limitations of existing measures of psychotic symptoms mean that a more comprehensive instrument, that captures the phenomenological overlap between trauma-related and hallucinatory experiences, is needed. This study details the development of the Questionnaire of Hallucinatory Experiences (QHE), a self-report measure that assesses the presence, frequency, and characteristics, of hallucinatory phenomena across 7 modalities, and their relation to traumatic experiences. An initial pilot study, based on the completion of the QHE by 44 individuals with recent experience of hallucinations, was conducted to evaluate the measure's psychometric properties against predetermined "go/revise/stop" criteria designed by the authors. Hallucinations, across sensory domains, were frequently reported by participants. Distress levels in relation to such experiences were high and impacted functioning, work, and employment. Participants agreed that hallucinations were associated with past trauma, however weak-moderate correlations between phenomenology and corresponding trauma subscales were found. Moderate internal consistency was reported (Cronbach's α=.64-.97) for subscales, indicating that all results are in line with "revise" criteria. Further iterations of the QHE are required with a focus on incorporating theoretically informed models of trauma in psychosis as well as increased service-user input to assess the relevance, comprehensive, and comprehensibility of items. It is hoped that the QHE will inform research into the mechanisms by which trauma impacts psychotic experiences and support collaborative formulations of hallucinations and their link with trauma and victimisation.

Keywords: hallucinations, trauma, psychosis, measures, assessment, posttraumatic stress, intrusions.

Introduction

Hallucinatory experiences, sensory perceptions in the absence of stimuli accessible to others (Bell et al., 2010), have traditionally been investigated in the context of schizophrenia-spectrum disorders. As a result, the focus of much of the literature has been on the phenomenological characteristics of hallucinations in psychosis (McCarthy-Jones et al., 2014), especially regarding those in the most reported modality, auditory hallucinations (AH) (McCarthy-Jones et al., 2017). More recently, the psychosis continuum theory (Johns & van Os, 2001) posited that the positive symptoms of psychosis, namely hallucinations and delusions, were evident across conditions, and in people in the general population without the associated loss of contact with reality which would warrant the need for treatment (Waters et al., 2018). This has precipitated the development of both transdiagnostic measures of psychotic phenomena (e.g., The Questionnaire for Psychotic Experiences, QPE (Rossell et al., 2019a), and unusual perceptual, sensory or anomalous experiences, including The Cardiff Anomalous Perceptions Scale (Bell et al., 2006); The Launay Slade Hallucination Scale (LSHS) revisions, (e.g., The Revised Hallucination Scale, (Morrison et al., 2000); LSHS-Modified (Larøi et al., 2004); LSHS-Modified-II (Larøi & Van der Linden, 2005), the Oxford-Liverpool Inventory of Feelings and Experiences, O-LIFE (Mason et al., 2005) and new scales, the Multi-Modality Unusual Sensory Experiences Questionnaire, MUSEQ (Mitchell et al., 2017) and The Transpersonal Experiences Questionnaire, TEQ (Heriot-Maitland et al., 2023), to be used in community samples.

Despite these impressive advances in the study of hallucinatory experiences, important limitations exist in the assessment of such phenomena (Rossell et al., 2019a). One significant weakness of many measures is the lack of clear, specific definitions of the hallucinatory constructs assessed. In the last decade, understandings of the psychosis continuum have expanded rapidly. Although advantageous, this has resulted in varied definitions for the range of experiences and symptoms found across this continuum and no standardisation of terms has been given (Seiler et al., 2020). This ambiguity is evident in the inclusion of other perceptual experiences, such as misperceptions (e.g., of images or

sounds) or illusions, in many hallucination or unusual sensory experiences (USE) measures (e.g., MUSEQ, RHS, QPE). The use of inconsistent definitions impedes the ability to accurately explore hallucinatory phenomenology and their prevalence (Aynsworth et al., 2017). The standardisation of psychotic symptom terminology is crucial to promote clear communication among researchers, clinicians, and the public, improve understanding of hallucinations and inform psychological treatment to those distressed by such experiences.

A second major limitation of most existing measures is their assessment of only a narrow range of hallucinatory modalities. Many available measures examine single hallucinatory modalities in psychosis (e.g., AHs; e.g., Psychotic Symptom Ratings Scales, PSYRATS (Haddock et al., 1999a)) or only assess such phenomena using global scores rather than independently rating hallucinatory modalities (e.g., the Positive and Negative Syndrome Scale, PANSS (Kay et al., 1987). Although recent versions of the LSHS (e.g., LSHS-M-II) explore hallucinations in several domains (auditory, visual, olfactory, tactile, and sensed presence), only one or two items per modality are included. The MUSEQ offers a more thorough measurement of 6 hallucinatory domains, incorporating 8 gustatory items. The measure is, however, limited by its inclusion of only hallucination and USE frequency ratings. Other important severity information is therefore missed, such as duration, distress, and impact. The QPE provides the most comprehensive assessment of hallucinations in all modalities bar the less commonly reported gustatory domain (Linszen et al., 2022). This measure also incorporates an item assessing the presence of multi-modal hallucinations (MMH), sensory experiences that occur simultaneously in two or more modalities (Toh et al., 2022). These hallucinations, as well as those in the tactile and olfactory modalities, are grouped into the “Hallucinations in other modalities” subscale, thus limiting the number of items specific to each type of sensory experience. Given the existence of hallucinations in varied sensory domains and recent findings that MMH are more prevalent in both psychiatric and organic disorders than once assumed (Montagnese et al., 2020), comprehensive and sensitive scales that broadly assess these phenomena are required. The exploration of MMHs in hallucination measures is especially pertinent given that such

experiences are associated with higher levels of psychological distress and adverse mental health outcomes (Laloyaux et al., 2019).

A related concern is the absence of measures investigating the phenomenological commonality between hallucinatory experiences and trauma reliving symptoms (e.g., flashbacks and intrusive memories). The relationship between trauma, in particular childhood mistreatment, and psychosis, is well established; several reviews and meta-analyses indicate that developmental trauma is associated with an increased risk of developing psychotic symptoms (Alameda et al., 2021; Bailey et al., 2018; Beards et al., 2013; Bendall et al., 2008; Kraan et al., 2015; Read et al., 2005) and more severe hallucinations and delusions in those who already have psychosis (Beards et al., 2013; Varese et al., 2012). A specific association between interpersonal abuse and voice hearing has been proposed (Hardy, 2017; van den Berg et al., 2023). Although often limited by small sample sizes, a growing body of research has explored such links; Peach et al (2020) reported that 33% (*n*=36) of their sample of young people with first episode psychosis experienced hallucinations with content directly related to past trauma and 67% described their hallucinations as thematically related to abusive experiences. Similarly, 68% of individuals identified their voice as either an abusive family member or other perpetrator in Corstens and Longden's (2013) data synthesis of 100 clinical cases. In a recent, more methodologically robust study, Van den Berg, and colleagues (2023) developed a structured coding frame to assess both independent and dependent links between trauma and auditory hallucinations. Relational and identity links were frequently reported by participants (*n*=73) (75%, 51%) and researchers (64%, 22%). Moreover, themes of cognitive-affective psychological threat, including feelings of powerlessness, shame, and self-blame characterised both the voices participants reported and their abuse histories. Negative self-beliefs in relation to trauma also significantly increased the likelihood of experiencing abusive, critical voices.

Such findings demonstrate the urgent need for accurate assessment of hallucinatory phenomenology and its overlap with trauma symptoms. Yet, to date, there are no existing,

validated, measures designed to explore the trauma to psychosis continuum of hallucinatory experiences. This is concerning given the high incidence of trauma among individuals reporting psychotic symptoms and the significant under-detection of post-traumatic stress in this population (Corstens & Longden, 2013). Therefore, the purpose of the current study was to develop a comprehensive, transdiagnostic measure of hallucinatory phenomena in a broad range of modalities (auditory, visual, olfactory, tactile, gustatory, sensed presence and MMH). The Questionnaire of Hallucinatory Experiences (QHE) assesses hallucinatory presence, phenomenology, and relation to traumatic experiences and is designed as a self-report measure to be used in clinical, non-clinical, and research contexts. The current study reports the development of the measure and preliminary pilot examination of its reliability.

Methods

Design

A three-phased, cross-sectional design was employed to develop a self-report measurement tool of hallucinatory phenomena (see Figure 2.1). Originally, a factor analytic study was planned for the third phase to inform item selection for the final measure. However, recruitment difficulties deemed this initial aim unviable. An internal pilot of the measure was therefore conducted to clarify the next development phases. The original research proposal, and revised analysis plan, are available via the links in Appendix 2.1 and 2.2. The study was approved by the University of Glasgow's Medical, Veterinary, Life Sciences (MVLS) Ethics Committee (ref: 200210004), see appendix 2.3.

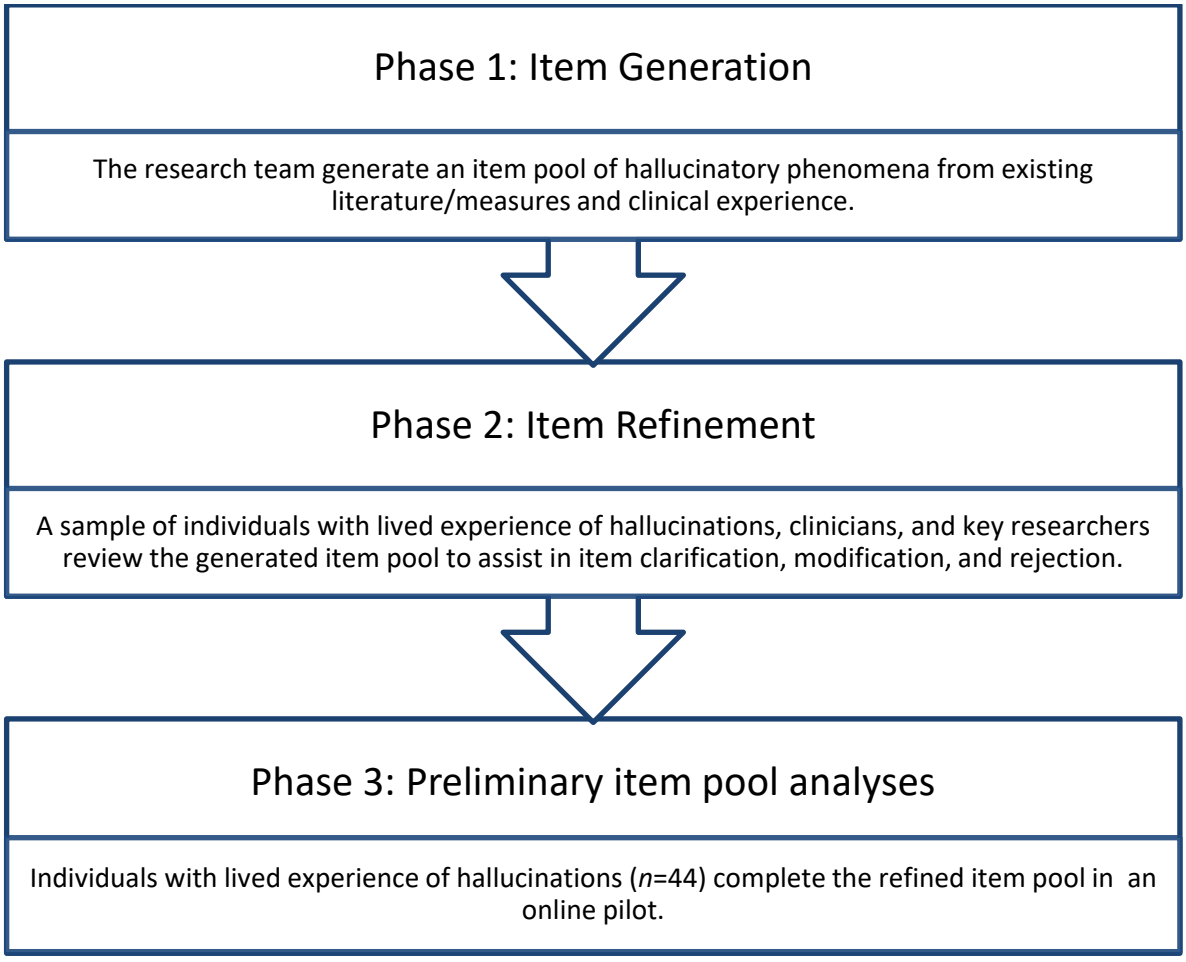

Figure 2. 1: Study Design Phases

Scale Development

Phase 1: Item Generation

A large, varied item pool was generated and refined in several phases. The objective of the item generation phase was to generate an exhaustive list of hallucinatory phenomena, across several modalities, and incorporate the perspectives of clinicians, researchers, and individuals with lived experience of hallucinations.

A broad range of existing theory, clinical experience and knowledge was utilised by the research team to generate an initial item pool. Potential items were designed following an extensive review of scales and interviews that assess hallucinations, unusual sensory experiences, perceptual anomalies, and trauma (PTSD/CPTSD) in different clinical groups (see Table 2.1). Measures from other literatures were also consulted in the creation of phenomenology (e.g., frequency, impact, control) ratings.

Phase 2: Item Refinement

Fifteen key researchers, individuals with lived experience of hallucinations, and clinician-academics in the psychosis, hallucinatory experiences, and trauma field were invited to review the initial item pool (*n*=184). Eight of the 15 individuals approached gave feedback on the relevance, comprehensiveness (i.e., whether key concepts were missing), and comprehensibility (i.e., clarity and wording of items and response options) of the items (Terwee et al., 2018). Following such feedback, revisions were made to the wording and length of questions, and the item pool was significantly decreased for ease of completion (*n*=151).

Table 2. 1: Scales and assessment measures reviewed during development of the QHE.

Self-report scales
Cardiff Anomalous Perception Scale (Bell et al., 2006)
Launay Slade Hallucination Scale (Launay & Slade, 1981)
Launay Slade Hallucination Scale-Modified (Larøi et al., 2004)
Launay Slade Hallucination Scale-Modified-II (Laroi & Van der Linden, 2005)
Oxford-Liverpool Inventory of Feelings and Experiences (Mason et al., 1995; Mason et al., 2005)
Multi-Modality Unusual Sensory Experiences Questionnaire (Mitchell et al., 2017)
The Transpersonal Experiences Questionnaire (Heriot-Maitland et al., 2023)
Revised Hallucination Scale (Morrison et al., 2000)
Clinical interview schedules
Psychotic Symptom Rating Scales (Haddock et al., 1999)
Questionnaire for Psychotic Experiences (Rossell et al., 2019a)
Queen Square Visual Hallucination Inventory (Williams et al., 2008)
North-East Visual Hallucinations Interview (Mosimann et al., 2008)
Scale for the Assessment of Positive Symptoms (Andreasen et al., 1991)
University of Miami Parkinson's Disease Hallucination Questionnaire (Papapetropoulos et al., 2008)
Neuropsychiatric Inventory (Cummings et al., 1994)
The Psycho-Sensory Hallucinations Scale (de Chazeron et al., 2015)
Positive and Negative Syndrome Scale (Kay et al., 1987)

The items in the revised QHE are arranged into 7 modules or subscales that correspond to the different hallucinatory domains: auditory hallucinations (AH), visual hallucinations (VH), gustatory hallucinations (GH), tactile hallucinations (TH), multi-modal hallucinations (MMH) and sensed presence (SP; see Figure 2.2). Each module begins with a screening question that asks about the presence of experiences within that hallucinatory domain in the last year. If endorsed, the module is included in the assessment. If the screening question is negative, the respondent is prompted to skip the questions from that module. However, if this screening question is endorsed, the respondent is directed to complete the remaining items in this module. Beyond presence, the QHE assesses other phenomenological aspects of such experiences: frequency, content, distress, control, impact, and beliefs about causation, as well as their association with trauma and adversity. 5-point Likert rating scales (anchored by 1= "*not at all",* 5 ="extremely") are used to assess phenomenological characteristics.

The trauma subscales contained between 7-10 items aimed to capture trauma-hallucination associations. Items assessed whether participants perceived a link with their hallucinations, in different modalities and past events, and if these triggered trauma reliving symptoms (e.g., flashbacks, intrusive memories, or nightmares) of past events. Four other items assessed potential relational, identity, and content links between hallucinations and past events (e.g., if auditory hallucinations sounded like/belonged to people involved in past traumatic events). Another item assessed if intrusive memories triggered hallucinations. Lastly, independent links between hallucinations and traumatic events (e.g., coping, and emotional responses) were included. The 40 items included in these subscales were rated on a 5-point Likert scale, according to participants' level of agreement with each statement. A further 14 open-ended questions are included to capture the content of hallucinatory experiences across the modalities, and the respondent's beliefs around the causation of their hallucinations.

Scoring

As stated, Likert scales were scored from 1 to 5, with higher scores indicating greater agreement and thus more significant difficulties/poorer quality of life.

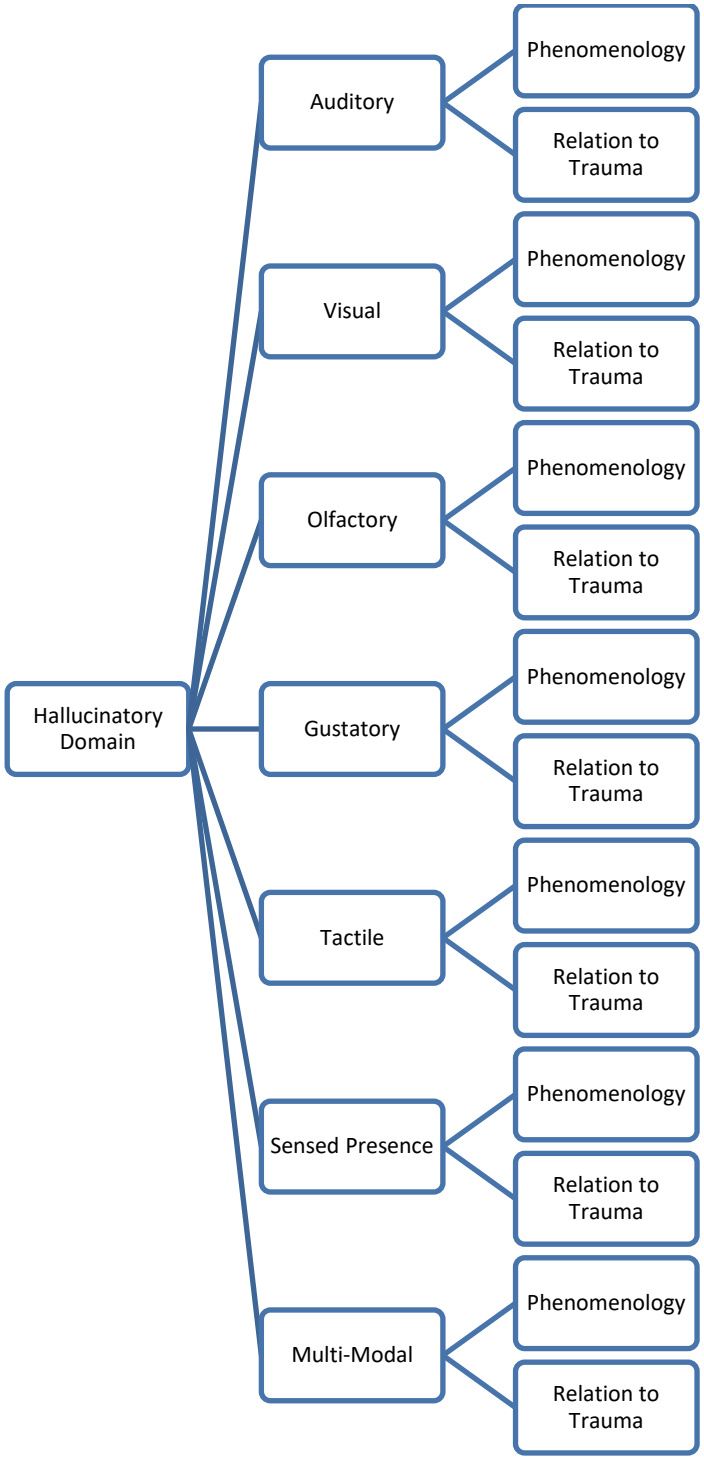

Figure 2. 2: QHE modules and structure

Phase 3: Preliminary item pool analysis

Participants

Participants were recruited online, via social media sites (Twitter, Instagram, Facebook, Redit), to take part in the study. Inclusion criteria for all participants were that they needed to be at least 18 years old, fluent English language speakers, have experienced auditory, visual, olfactory, gustatory, tactile, multi-modal hallucinations or a sensed presence in the last year and have access to a computer or mobile electronic device.

Procedure

A link to the QHE (see appendix 2.4) was given in all social media posts and advertisements. Individuals were initially directed to a Participant Information Sheet (PIS) with details of the study to assist them to decide if they wished to take part. Those who opted to participate provided informed consent and were then directed to complete demographic information, e.g., their age in years, ethnicity, and sex. Following this, participants completed the QHE and were invited to provide their email address if they wished to be sent a copy of the results once the study was completed. A copy of the consent form and PIS are available via the links in 2.5 and 2.6.

Data Analysis

Existing literature and theory on scale development (Clark & Watson, 2016; Clark & Watson, 2019; Wood & Boyce, 2017) was employed to define novel Proceed, Revise and Stop criteria (see Table 2.2) to assess the psychometric properties of the QHE and clarify the next development phases. Such criteria were detailed in the revised analysis plan which was approved by the University of Glasgow's MVLS ethics committee before data collection ceased.

Table 2. 2: Go/Revise/Stop Criteria for data analysis

Go: *No revision required to the QHE ahead of factorial analytic study*.	**Revise:** *Further iteration to the existing item pool is required.*	**Stop:** *Substantial revision required via further stakeholder consultation and qualitative content validity studies.*
Variability: A range of responses is reported for all subscale items, indicating that they are useful and varied. Spearman Rank Order correlations, (with 95% CIs) within each hallucinatory domain demonstrate adequate variance; no items are highly correlated (rho=<0.7+).	Variability: Various subscale items are highly skewed with unbalanced distributions. Spearman Rank Order correlations, (with 95% Cis) within each hallucinatory domain report high correlations between items (rho=>0.7+). Caution will be taken when reporting item variability, however, given the likelihood that the small sample size may produce unstable correlational results.	Variability: Most items show skewed and unbalanced response distributions. Spearman Rank Order correlations, (with 95% CIs) within each hallucinatory domain, indicate that items are highly correlated with each other (rho=>0.8+).
Reliability: The QHE possess good internal consistency with Cronbach's alpha coefficients of 0.80 or above for all subscales. Mean inter-item correlations fall within the intended range of 0.15-0.50, suggesting that while items are reasonably homogenous, they contain sufficient variance.	Reliability: 95% Confidence internals indicate that reliability estimates are not possible, or the QHE is found to possess moderate internal consistency, with Cronbach's alpha coefficients of 0.60 or below for several subscales. Some mean inter-item correlations fall out with the intended range.	Reliability: Confidence internals indicate that reliability estimations are not possible, or the QHE has low internal consistency (α=0.30) on several/all the subscales. Most mean inter-item correlations fall out with the intended range.
Associations between hallucinatory modality and traumatic experiences subscales: Spearman Rank Order correlations (with 95% CIs) demonstrate strong, positive correlations between hallucinatory domain and the traumatic experiences subscales, with effect sizes of 0.50 or above.	Associations between hallucinatory domain and traumatic experiences subscales: Spearman Rank Order correlations (with 95% CIs) report small to moderate effect sizes of 0.30 or above.	Associations between hallucinatory domain and traumatic experiences subscales: Spearman Rank Order correlations (with 95% CIs) report small correlations between these subscales, with effect sizes of 0.10 or below.
Qualitative Items: Content analysis does not generate novel constructs, ideas, or items that would indicate that further revision to the measure is required.	Qualitative Items: Content analysis highlights some novel constructs, ideas, or items which suggest that further iteration to the existing item pool is warranted.	Qualitative Items: Content analysis reports many new constructs or ideas, indicating that significant item generation and refinement is required.

The association between each hallucinatory modality and the respective trauma subscale were investigated by Spearman Rank Order correlations (with 95% confidence intervals) Internal consistency of subscales was determined using Cronbach's α coefficient. Although many contemporary researchers characterise reliabilities of .60 or .70 as good(Clark & Watson, 2019),Cronbach's α values of above .80 were set in the "Proceed/Revise/Stop criteria" as these indicate equivalence and ensure that the items are not measuring too disparate concepts (Terwee et al., 2007).

A content analysis of the 14 qualitative items (content of hallucinations and beliefs about causation) was conducted. Keywords in participant responses were initially coded and used to generate themes.

Results

Participants

A total of 44 participants (female *n*=37; male *n*=6; other *n*=1) completed the QHE. The inclusion criteria stipulated that individuals must have experienced hallucinations in the past year to participate. However, the questionnaire was worded in such a way that in practice, only those who had experienced hallucinations in the last month were eligible to complete the QHE. Thus, 5 individuals endorsed the screening item in relation to the experience of hallucinations in the last year but could not participate as they did not report such experiences in the last month. Basic participant characteristics are presented in table 2.3. Most of the sample (n=37, 84.1%) were female, over half (n=25, 56.8%) were aged between 25-44 years old and thirty-eight (86.4%) participants were white.

Table 2. 3: Demographic Characteristics of Participants

Variable	***n* (%)**
Gender	
Male	6 (13.6)
Female	37 (84.1)
Other	1 (2.3)
Age (years)	
18-24	3 (6.8)
25-34	13 (29.5)
35-44	12 (27.3)
45-54	10 (22.7)
55-64	4 (9.1)
65+	2 (4.5)
Ethnicity	
Asian	1 (2.3)
Mixed/Multiple Ethnicities	4 (9.1)
Other Ethnic Group	1 (2.3)
White	38 (86.4)

Figure 2.3 shows the range of hallucinatory modalities endorsed by participants. Most of the sample (*n*=33, 75%) had experienced auditory hallucinations in the past month. Thirty

(68.2%) reported visual hallucinations in this time frame and 23 (52.3%) indicated that they had felt a presence in the last month. Participants reported the experience of tactile hallucinations on their body (n=20, 45.5%) and inside their body (n=12, 27.3%). Multi-modal hallucinations were experienced by 19 (43.2%) individuals. Unsurprisingly, hallucinations in the olfactory and gustatory modalities were less commonly reported; n=10, (22.7%) for OH and 3 (6.8%) for gustatory hallucinations. Given these small numbers for OH and GH (n=3 and 10 respectively), we did not perform internal consistency analyses for these subscales.

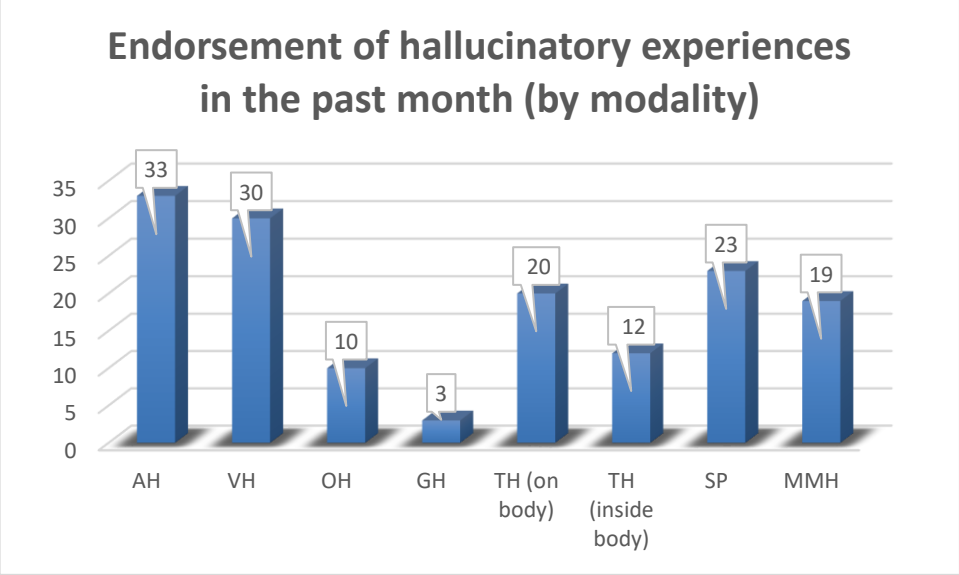

Figure 2. 3: Number of participants reporting the experience of hallucinations (by modality) in the past month.

Note: AH=Auditory Hallucinations; VH=Visual Hallucinations; OH=Olfactory Hallucinations; GH=Gustatory Hallucinations; TH=Tactile Hallucinations; SP=Sensed Presence; MMH=Multi-Modal Hallucinations.

Acceptability

After consent was obtained, 44 (83%) of the individuals who began the questionnaire finished the QHE, demonstrating high completion rates. Four individuals (7%) began the QHE but did not finish it. A further 5 (10%) people began the questionnaire but were excluded as they did not report experiencing hallucinations in the last month. The average time to finish the questionnaire was 18.5 minutes (range 3.2 to 81 minutes), depending on the number of hallucinatory modalities endorsed and response speed. One outlier of a completion time of 998 minutes was found.

Subscale Descriptive Statistics

Tables 2.4-2.8 present descriptive statistics of participants' responses to the QHE phenomenology subscale items for the auditory, visual, tactile, multi-modal hallucination and sensed presence modalities. Olfactory and gustatory descriptive statistics are not reported due to the very small sample size of each of these subgroups. Participants reported experiencing hallucinations frequently; auditory hallucinations occurred most often (scored from 0, less than once a week, to 4, at least once an hour) (*Median*=3.00), with most individuals reporting hearing things that others could not at least once every two hours (see table 2.4). Participants reported visual (*Median*=2:00), tactile (*Median*=2.00) and multi-modal hallucinations (*Median*=2:00) and sensed presence (*Median*=2.00) as occurring at least once a day (see tables 2.5-8).

Table 2.4: Subscale Internal Consistency–Auditory Hallucination (AH) Phenomenology Scale

Subscale	Item	Mean	SD	Median	Range	N	Conbach's α	95% CI
Phenomenology								
	Presence: During the past month, have you heard sounds that other people could not hear?	1.30	0.55	1.00	2	44		
	Frequency: During the past month, how often have you heard things that other people could not hear?	3.03	1.38	3.00	4	33		
	Duration: How long do you usually hear these things for?	2.94	1.35	3.00	4	33		
	Loudness: How loud are the things you hear?	2.06	0.79	2.00	3	33		
	Distress: How much do the things you hear distress you?	2.88	1.11	3.00	4	33		
	Negative content: How negative is the content of the things you hear?	2.97	1.26	3.00	4	33		
	Control: To what extent are you able to control the things you hear and dismiss them at will?	2.18	1.10	2.00	4	33		
	Impact on functioning: The things I hear negatively impact my day-to-day functioning?	2.82	1.19	3.00	4	33		
	Impact on work/study: The things I hear negatively impact my ability to work or study.	3.39	1.37	3.00	5	33		
	Impact on Relationships: The things I hear negatively impact my ability to form and maintain relationships with others.	2.91	1.21	3.00	4	33		
	All items	**2.62**	**0.69**	**2.50**	**2.70**	**33**	**0.80**	**0.68 - 0.89**

Table 2.5: Subscale Internal Consistency – Visual Hallucination (VH) Phenomenology Scale

Subscale	Item	Mean	SD	Median	Range	N	Conbach's α	95% CI
Phenomenology								
	Presence : During the past month, have you seen anything (e.g., objects, visual patterns, people etc) that other people could not?	1.34	0.53	1.00	2.00	44		
	Frequency: During the past month, how often have you seen things that other people could not hear?	2.37	0.96	2.00	4.00	30		
	Duration : How long do you usually see these things for?	2.03	1.33	1.50	4.00	30		
	Clarity : How clear are the things you see?	2.73	0.87	3.00	3.00	30		
	Distress: How much do these sights distress you?	2.77	1.17	3.00	4.00	30		
	Control : To what extent are you able to control the things you see and dismiss them at will?	2.00	1.20	1.50	4.00	30		
	Impact on functioning : The things I see negatively impact my day-to-day functioning?	2.40	1.13	2.00	4.00	30		
	Impact on work/study: The things I see negatively impact my ability to work or study.	2.83	1.39	3.00	5.00	30		
	Impact on Relationships: The things I see negatively impact my ability to form and maintain relationships with others.	2.30	1.06	2.00	4.00	30		
	All items	**2.27**	**0.55**	**2.17**	**2.22**	**30**	**0.64**	**0.41 - 0.81**

Table 2.6: Subscale Internal Consistency-Tactile Hallucination (TH) Phenomenology Scale

Subscale	Item	Mean	SD	Median	Range	N	Conbach's α	95% CI
Phenomenology								
	Presence of TH on body: During the past month, have you expereinced any sensations on your body (e.g., burning, tingling, scraping, heat) despite being aware that there was nothing touching you?	1.61	0.62	2.00	2.00	44.00		
	Presence of TH inside body: During the past month, have you felt sensations inside your body (e.g., pulling movement, burning) that you could not account for?	1.80	0.55	2.00	2.00	44.00		
	Frequency: During the past month, how often have you experienced sensations on or inside your body that you could not explain?	2.00	1.06	2.00	4.00	24.00		
	Duration: How long do you usually experience these sensations for?	2.38	1.28	2.00	4.00	24.00		
	Intensity: How intense are the sensations you feel?	2.38	0.87	3.00	3.00	24.00		
	Distress: How much do these sensations distress you?	3.71	1.20	4.00	4.00	24.00		
	Control: To what extent are you able to control these sensations and dismiss them at will?	1.88	1.42	3.00	2.00	24.00		
	Impact on functioning: The sensations I feel negatively impact my day-to-day functioning?	2.75	1.29	3.00	4.00	24.00		
	Impact on work/study: The things I feel negatively impact my ability to work or study.	2.88	1.42	3.00	4.00	24.00		
	Impact on Relationships: The things I feel negatively impact my ability to form and maintain relationships with others.	2.46	1.29	2.00	4.00	24.00		
	All Items	**2.37**	**0.57**	**2.50**	**2.10**	**24.00**	**0.69**	**0.46-0.84**

Table 2. 7: Subscale Internal Consistency- Sensed Presence (SP) Hallucination Phenomenology Subscale

Subscale	Item	Mean	SD	Median	Range	N	Conbach's α	95% CI
Phenomenology								
	Presence : During the past month, have you felt a presence of someone (e.g., behind/beside you, or in another room) even though you could not see, hear, touch, or smell them?	1.55	0.63	1.00	2.00	44.00		
	Frequency: During the past month, how often have you felt a presence of someone even though you could not see, hear, touch, or smell them?	2.48	1.16	2.00	4.00	23.00		
	Duration: How long do you usually feel this presence for?	2.78	1.35	3.00	4.00	23.00		
	Intensity: How intensely do you feel this presence?	2.91	0.79	3.00	4.00	23.00		
	Distress: How much does this presence distress you?	3.00	1.31	3.00	4.00	23.00		
	Control: To what extent are you able to control this presence and dismiss it at will?	1.65	0.89	1.00	3.00	23.00		
	Impact on functioning: These experience negatively impact my day-to-day functioning.	2.22	1.00	2.00	3.00	23.00		
	Impact on work/study: These experience negatively impact my ability to work or study.	2.65	1.58	3.00	5.00	23.00		
	Impact on Relationships: These experience negatively impact my ability to form and maintain relationships with others.	2.30	1.30	2.00	4.00	23.00		
	All Items	**3.02**	**0.88**	**3.00**	**3.13**	**23.00**	**0.65**	**0.39-0.85**

Table 2.8: Subscale Internal Consistency- Multi-Modal (MMH) Hallucination Phenomenology Subscale

Subscale	Item	Mean	SD	Median	Range	N	Conbach's α	95% CI
Phenomenology								
	Presence: During the past month, have you had experiences which involved more than one of your senses at the same time (e.g., seeing and hearing someone that other people do not see, smelling, tasting, or hearing things that no-one else can)?	1.64	0.61	2.00	2.00	44		
	Frequency: During the past month, how often have you had experienced which involve more than one sense at the same time?	1.95	0.91	2.00	3.00	19		
	Duration: How long do you usually last for?	2.26	0.87	2.00	3.00	19		
	Intensity: How intense are these experiences?	2.79	0.86	3.00	2.00	19		
	Distress: How much do these experiences distress you?	3.58	1.31	3.00	3.00	19		
	Control: To what extent are you able to control these experiences and dismiss them at will?	1.84	1.02	2.00	3.00	19		
	Impact on functioning: These experience negatively impact my day-to-day functioning.	2.95	1.18	3.00	4.00	19		
	Impact on work/study: These experience negatively impact my ability to work or study.	3.47	1.43	4.00	4.00	19		
	Impact on Relationships: These experience negatively impact my ability to form and maintain relationships with others.	2.84	1.17	3.00	4.00	19		
	All Items	**2.52**	**0.58**	**2.33**	**2.11**	**19**	**0.72**	**0.48-0.88**

Auditory, multi-modal, and sensed presence hallucinations were reported to last for several hours (scored from 0=a few seconds, to 4=almost continuously, *Median*=3.00). Tactile (*Median*=2.00) and visual (*Median*=1.50) hallucinations lasted from a few minutes to up to an hour.

Participants reported high levels of distress in relation to their hallucinatory experiences and little control over such phenomena. Impact ratings (scored from 0=not at all, to 4=extremely) were also high across hallucinatory domains.

Relation to Traumatic Experiences Subscales

Tables 2.9-2.13 present descriptive statistics of participants' responses to relation to traumatic experiences subscale items for the auditory, visual, tactile, multi-modal hallucination and sensed presence modalities. Thirty-two (80%) of items in the trauma subscales received a rating of 3 (*Median*=3.00), indicating that participants agreed with the items assessing the hallucinations-trauma link.

Auditory Hallucination (AH) and corresponding trauma subscale:

The median rating (*Median*=3.00), for all 10 trauma items in the AH subscale indicated participant agreement that their AHs were related to trauma-reliving symptoms, see table 2.9.

Table 2.9: Subscale Internal Consistency – Auditory Hallucination (AH) Relation to Traumatic Experiences Scale

Subscale	Item	Mean	SD	Median	Range	N	Conbach's α	95% CI
Relation to Traumatic Experiences								
	The things I hear are related to past events.	3.30	1.05	3.00	4.0	33		
	The things I hear trigger powerful images of past events which feel like they are happening again in the here and now.	2.97	1.43	3.00	4.0	33		
	The things I hear trigger vivid memories of a past event(s).	3.00	1.25	3.00	4.0	33		
	Certain memories of past events lead me to hear things that other people do not.	3.24	1.25	3.00	4.0	33		
	The things I hear lead me to have upsetting dreams which are related to, or about, past events.	3.18	1.31	3.00	4.0	33		
	I cope with the things I hear in a similar way to how I coped with the past event(s).	3.18	1.13	3.00	4.0	33		
	The things I hear make me feel the same as how I felt during the past event(s).	3.03	1.21	3.00	4.0	33		
	The things I hear sound like, or belong to, people involved in the past event(s).	2.97	1.36	3.00	4.0	33		
	The things I hear are the same things that were said to me during the past event(s).	2.88	1.32	3.00	4.0	33		
	The things I hear treat me, or behave similarly to, the way people treated me during the past event(s).	3.15	1.35	3.00	4.0	33		
	All Items	**3.09**	**0.98**	**3.10**	**4.00**	**33**	**0.93**	**0.88 - 0.96**

Visual Hallucination (VH) and corresponding trauma subscale:

Table 2.10 illustrates that participants agreed that their VHs were linked to traumatic memories and nightmares (*Median*s=3.00). Intrusive memories were found to lead to VHs (*Median*=3.00). Behavioural and emotional response links between VH and past trauma, (*Median*=3.00) were reported.

Table 2.10: Subscale Internal Consistency – Visual Hallucination (VH) Relation to Traumatic Experience Scale

Subscale	Item	Mean	SD	Median	Range	N	Conbach's α	95% CI
Relation to Traumatic Experiences								
	The things I see are related to past events.	2.73	1.53	2.00	4.00	30		
	The things I see trigger powerful images of past events which feel like they are happening again in the here and now.	2.37	1.10	2.00	4.00	30		
	The things I see trigger vivid memories of a past event(s).	2.77	1.33	3.00	4.00	30		
	Certain memories of past events lead me to see things that other people do not.	2.80	1.27	3.00	4.00	30		
	The things I see lead me to have upsetting dreams which are related to, or about, past events.	3.83	1.39	3.00	4.00	30		
	I cope with the things I see in a similar way to how I coped with the past event(s).	3.83	1.29	3.00	4.00	30		
	The things I see make me feel the same as how I felt during the past event(s).	2.83	1.26	3.00	4.00	30		
	All Items	**2.74**	**1.06**	**2.57**	**4.00**	**30**	**0.91**	**0.85 - 0.95**

Tactile Hallucination (TH) and corresponding trauma subscale:

Median ratings for 5 of the 7 (71.4%) TH trauma subscale items were 3 (see table 2.11), indicating that participants agreed that the things they felt were related to trauma and triggered intrusive memories.

Table 2.11: Subscale Internal Consistency-Tactile Hallucination (TH) Relation to Traumatic Experiences Scale

Subscale	Item	Mean	SD	Median	Range	N	Conbach's α	95% CI
Relation to Traumatic Experiences								
	The sensations I feel are related to past events.	3.00	1.50	3.00	4.00	24.00		
	The sensations I feel trigger powerful images of past events which feel like they are happening again in the here and now.	2.75	1.48	2.00	4.00	24.00		
	The sensations I feel trigger vivid memories of a past event(s).	2.71	1.40	2.00	4.00	24.00		
	Certain memories of past events lead me to suddenly experience feelings on or inside my body when there is nothing there.	3.04	1.63	3.00	4.00	24.00		
	The sensations I feel lead me to have upsetting dreams which are related to, or about, past events.	2.92	1.50	3.00	4.00	24.00		
	I cope with the sensations I feel in a similar way to how I coped with the past event(s).	2.75	1.33	3.00	4.00	24.00		
	The sensations I feel make me feel the same as how I felt during the past event(s).	2.88	1.42	3.00	4.00	24.00		
	All Items	**2.86**	**1.32**	**2.86**	**4.00**	**24.00**	**0.96**	**0.93-0.98**

Sensed Presence (SP) and corresponding trauma subscale:

The median rating assessing the relational link between SP and trauma was 4 (see table 2.12); participants strongly agreed that the presence they felt was related to trauma and thought that the SP was an individual involved in their past (*Median*=3.00). Independent links (e.g., coping, and emotional response) between SP and trauma were endorsed (*Median*=3.00), as were SP and reliving symptoms (flashbacks and nightmares, *Median*=3.00).

Table 2.12: Subscale Internal Consistency- Sensed Presence (SP) Hallucination Relation to Traumatic Experiences Subscale

Subscale	Item	Mean	SD	Median	Range	N	Conbach's α	95% CI
Relation to Traumatic Experiences								
	These multi-sensory experiences are related to past events.	3.52	1.12	4.00	3.00	23.00		
	These multi-sensory experiences trigger powerful images of past events which feel like they are happening again in the here and now.	2.87	1.29	3.00	4.00	23.00		
	These multi-sensory experiences trigger vivid memories of a past event(s).	2.78	1.28	2.00	4.00	23.00		
	Certain memories of past events lead me to have these multi-sensory experiences.	2.96	1.22	2.00	4.00	23.00		
	These multi-sensory experiences lead me to have upsetting dreams which are related to, or about, past events.	2.74	1.10	3.00	4.00	23.00		
	I cope with these multi-sensory experiences in a similar way to how I coped with the past event(s).	2.91	1.04	3.00	4.00	23.00		
	These multi-sensory experiences make me feel the same as how I felt during the past event(s).	3.22	1.04	3.00	3.00	23.00		
	These multi-sensory experiences include a person involved in past event(s).	3.13	1.33	3.00	4.00	23.00		
	All Items	**3.02**	**0.88**	**3.00**	**3.13**	**23.00**	**0.89**	**0.80-0.95**

Multi-Modal Hallucination (MMH) and corresponding trauma subscale:

Higher rates of agreement were reported for the items in this subscale, as is illustrated in table 2.13 Participants strongly agreed (*Median*=4.00) that MMH were related to the past and triggered vivid memories (*Median*=3.32). Individuals agreed that MMH triggered reliving symptoms (*Median*=3.00).

Table 2.13: Subscale Internal Consistency- Multi-Modal (MMH) Hallucination Relation to Traumatic Experiences Subscale

Subscale	Item	Mean	SD	Median	Range	N	Conbach's α	95% CI
Relation to Traumatic Experiences								
	These multi-sensory experiences are related to past events.	3.26	1.49	4.00	4.00	19		
	These multi-sensory experiences trigger powerful images of past events which feel like they are happening again in the here and now.	3.11	1.49	3.00	4.00	19		
	These multi-sensory experiences trigger vivid memories of a past event(s).	3.32	1.34	3.32	4.00	19		
	Certain memories of past events lead me to have these multi-sensory experiences.	3.16	1.39	3.00	4.00	19		
	These multi-sensory experiences lead me to have upsetting dreams which are related to, or about, past events.	3.11	1.49	3.00	4.00	19		
	I cope with these multi-sensory experiences in a similar way to how I coped with the past event(s).	2.84	1.29	3.00	4.00	19		
	These multi-sensory experiences make me feel the same as how I felt during the past event(s).	3.32	1.29	3.00	4.00	19		
	These multi-sensory experiences include a person involved in past event(s).	3.05	1.58	3.00	4.00	19		
	All Items	**3.14**	**1.28**	**3.00**	**4.00**	**19**	**0.97**	**0.94-0.99**

Reliability:

Internal Consistency

Tables 2.4 -2.13also display the internal consistency of the auditory, visual, tactile, multi-modal hallucination and sensed presence subscales. The olfactory and gustatory subscales were excluded due to their small sample size. All QHE trauma subscales demonstrated excellent internal consistency with Cronbach's α ranging from .89 to .97, illustrating the high inter-relatedness of these subscale items.

The phenomenology subscale for auditory hallucinations (α=.80, 95%CI: 0.68-0.86) also showed good internal consistency. However, the visual (α =.64, 95%CI:0.41-0.81), tactile (α =0.69, 95% CI: 0.46-0.84) and multi-modal hallucination phenomenology subscale (α=0.72, 95%CI: 0.48-0.88) showed moderate internal consistency. A relatively wider 95% CIs for these subscales (compared to auditory hallucinations) were noted. This is likely to be a function of the smaller sample sizes in these domains. Moreover, Cronbach's alpha is grounded in tau-equivalent theory which assumes that each questionnaire item measures the same trait on the same scale(Tavakol & Dennick, 2011). Although the items in the phenomenology subscales were designed to test the construct of hallucination severity, items assessing frequency, duration, and control of such experiences likely reduce item interrelatedness. Cronbach alpha values are also sensitive to the number of items in the scale(DeVellis, 2016). Scales with fewer than ten items commonly violate the assumption of tau-equivalence and underestimate reliability(Tavakol & Dennick, 2011). Both the VH and MMH phenomenology subscales contained 9 items which likely contributes to the lower Cronbach values.

Association between hallucinatory modality and trauma subscales

Table 2.14 displays the Spearman rho correlation coefficients (with 95% confidence intervals) between hallucinatory modality and the respective trauma subscale.

A moderate, positive correlation was found between the visual and corresponding trauma scale (*rho*=.50, CI:0.16-0.73), with higher scores on the visual hallucinations being

associated with higher trauma score. However, the wide confidence interval suggests a lack of certainty regarding the magnitude of correlation observed. A moderate, positive correlation, with a wide confidence interval, was also observed between the tactile hallucination phenomenology and the trauma subscales (*rho*=.49, CI: 0.10-0.75). The auditory phenomenology and trauma subscales were weakly correlated (*rho*=.27, CI: -0.10-0.57). The confidence interval crossed zero, again demonstrating high levels of uncertainty regarding this signal. A strong, negative correlation was observed for the multi-modal hallucination phenomenology and trauma subscales (*rho*=-0.56, CI: -0.85- -0.03). No correlation was observed between the sensed presence phenomenology and trauma subscales (*rho*=0.00, CI: -0.42-0.43).

Table 2. 14: Associations between hallucinatory phenomenology and trauma subscales

Associations between hallucinatory modality and trauma subscales	**Spearman's *rho* (95% confidence interval)**
Auditory Phenomenology-AH trauma subscale	n=33 0.27 (-0.10-0.57)
Visual Phenomenology-VH trauma subscale	n=30 0.50 (0.16-0.73)
Sensed Presence Phenomenology-SP trauma subscale	n=23 0.00 (-0.42-0.43)
Tactile Phenomenology-TH trauma subscale	n=20 0.49 (0.10-0.75)
Multimodal Phenomenology-MMH trauma subscale	n=19 -0.56 (-0.85- -0.03)
Gustatory Phenomenology-GH trauma subscale	n=3 Not Reported
Olfactory Phenomenology-OH, trauma subscale	n=10 Not reported

Qualitative Items:

The numbers and percentages reported in the section below are calculated according to the total amount of participants who endorsed experiencing hallucinations in each modality.

Content of Auditory Hallucinations:

The content of auditory hallucinations reported by participants was coded into two broad themes: voices and other sounds. Of the 33 respondents who disclosed experiencing auditory hallucinations, most (*n*=28, 85%) reported hearing voices. These could be further grouped into voices known to the participants (e.g., one respondent reported hearing their parents calling their name, two individuals heard their abuser's voice and another described hearing various characters, some of which were ancestors), and unknown voices. Five (15%) individuals advised that the unknown voices passed comment on, or summarised, their activity. Indistinct chatter or conversations were also commonly reported (2, 6%). Four (12%) individuals described the voices as negative; one respondent reported that the voices

were *"demeaning and cruel*" and 3 (9%) other individuals advised that the voices screamed at them. Respondents reported hearing several sounds which others did not. These included children laughing, babies, or animals crying as well as the noise of music, sirens, phones ringing, thuds, crashes, and chimes.

Content of Visual Hallucinations:

Participants (n=30) reported a variety of complex and simple visual hallucinations. Complex visual hallucinations included people, figures, faces, objects, animals (e.g., cats, spiders, and bugs). Most respondents (28, 93%) did not detail whether the people they saw were known to them. However, one individual described the woman they saw as like their mother. Another 2 (6.7%) participants reported seeing things related to past trauma; one individual saw the face of their abuser in scenes that appeared in front of them and the other saw shadows when they were in the vicinity of the place, they were abused in. Three (10%) individuals also recorded seeing distressing images of dismembered body parts, blood, or dead people. Simple visual hallucinations included shadows, flashes of light or colour and were reported by 14 (46.7%) of participants who experienced hallucinations. Visual distortions (e.g., a fire hydrant turning into a dog, street signs changing, faces and eyes distorting) were also reported by 2 (6.7%) of this subgroup.

Content of Olfactory Hallucinations:

Most smells (n=5, 50%) reported by the 10 participants who experienced such hallucinations were negative or threat based, e.g., rotting food, garbage, cigarette smoke, urine, blood, and gas. Two (20%) individuals recorded smelling their abuser. Other smells reported included aftershave, deodorant, food cooking, herbs, and flowers.

Content of Tactile Hallucinations:

A variety of tactile hallucinations were reported by the 20 participants who disclosed such experiences. These were coded into sensations felt on and inside the body. Sensations reported on the body included tingling, heat, numbness, paralysis, and pain. Four (20%)

individuals reported feeling that insects, bugs, or fleas were crawling on their skin. Many participants (10, 50%) noted feeling that they had been touched, tapped, pinched, pushed, grabbed, or tickled by someone/thing. Internal sensations reported by 2 (10%) individuals were sexual activity including one participant who described re-experiencing rape. Other sensations included feeling "cellular activity", things inside their stomach, or under their skin.

Content of Sensed Presence:

Sensed presences, reported by 23 individuals, were coded into two themes; felt presence of someone known to the participant and supernatural entities. Two respondents (8.7%) advised feeling that a deceased family member or friend was with them. Other sensed presences included a romantic partner, an estranged parent, and an abuser. Supernatural or spiritual entities (e.g., ghosts, God, or an energy) were also reported by 10 (43.5%). Three (13%) respondents indicated that the entities were sinister or "*evil."*

Content of Multi-Modal Hallucinations:

Seven (35%) of the 19 respondents who reported multi-modal hallucinations advised that these occurred simultaneously within the visual and auditory hallucinations modalities. One participant detailed seeing and hearing their abuser and another reported reliving sexual abuse. Four (21.1%) individuals reported that tactile hallucinations accompanied those in the visual and auditory domains. Two (10.5%) responses detailed multi-modal hallucinations involving sensing a presence around them and feeling this entity touch their skin.

Participant beliefs about the causes of their hallucinations:

The 7 items on participants' beliefs about the causality of their hallucinations were coded and categorised into 5 themes. Out of the 90 participant responses recorded by participants detailing their understanding of their hallucinatory experiences, 17 (18.9%) viewed hallucinations as symptomatic of either a mental health diagnosis, namely schizophrenia, or bi-polar disorder, or neurodevelopmental disorder including Autism.

Most participant responses (33, 36.7%) believed traumatic experiences and abuse to be key contributors to the onset of hallucinations. Trauma-related understandings of hallucinations included having a diagnosis of Post-Traumatic and Complex-Post Traumatic Stress Disorder or viewing such experiences as dissociative episodes. Three (3.3%) participant responses described such experiences as "*trauma flashbacks".* One participant described the association between their hallucinations and previous trauma; they reported that engaging in sexual activity triggered hallucinations about their prior sexual abuse. Another individual reported that trauma memories precipitated their hallucinations and believed that these were evidence of punishment. Grief was also reported as a traumatic, adverse experience which resulted in hallucinations.

Related to this, 8 (8%) of participant responses detailed an association between the onset of hallucinations and periods of heightened stress. Physical health issues such as migraines and insomnia were also highlighted as perceived hallucinatory causes in 5 (5.6%) responses.

Table 2.15: Summary of Analyses According to Go/Revise/Stop Criteria

Preliminary Analyses	*Go/Revise/Stop Criteria*		
	Go: No revision required to the QHE ahead of factorial analytic study.	**Revise:** Further iteration to the existing item pool is required.	**Stop:** Substantial revision required via further stakeholder consultation and qualitative content validity studies.
Variability		Unbalanced distributions were noted in the phenomenology subscales. However, item variability was likely impacted by the small sample sizes.	
Reliability: Internal Consistency		4 hallucinatory phenomenology subscales (VH, TH, SP and MMH) reported only moderate internal consistency (Cronbach's α ranged from .64-.72).Wide 95% Confidence internals limit confidence in the certainty of these estimates.	
Associations between hallucinatory modality and traumatic experiences subscales			No correlation was observed between the sensed presence phenomenology and trauma subscales (rho=0.00, CI: -0.42-0.43). A strong, negative correlation was observed for the multi-modal hallucination phenomenology and trauma subscales (*rho* =-0.56, CI: -0.85- -0.03).
Qualitative Items	Participants' responses were in line with the constructs included in the QHE. No novel constructs, ideas, or items were reported that would indicate that further revision to the measure is required.		

Table 2.15 summarises the results of the performed analyses in relation to the Go/Revise/Stop criteria. Although a range of responses were reported for the items in

traumatic experiences subscales, unbalanced distributions were found in the phenomenology subscales. Item variability was therefore in line with the “revise” criteria, indicating that further iteration to the phenomenology item pools is required. Likewise, the QHE was found to possess moderate reliability, with 4 subscales (VH, TH, MMH, SP phenomenology subscales) reporting Cronbach alpha coefficients in line with the revise criteria (Cronbach’s α 64-.72). This suggests poor inter-relatedness between items or the presence of heterogeneous constructs in the phenomenology subscales. Spearman Rank Order correlations between the hallucinatory domain and traumatic experiences subscales revealed inconsistent findings. Large, positive correlations were found between the visual (*rho*=.50, CI:0.16-0.73) and tactile (*rho*=.49, CI: 0.10-0.75) phenomenology scales and their corresponding trauma scale. However, wide confidence intervals were noted. Spearman Rank Order correlations for the multi-modal (*rho*=-0.56, CI: -0.85- -0.03) and sensed presence (*rho*=0.00, CI: -0.42-0.43) scales and their corresponding trauma scales, were in a line with the stop criteria. However, content analysis of participants’ qualitative responses were in line with the constructs included in the QHE and so met the Go criteria. Overall, preliminary investigation of the QHE’s reliability met the “revise” criteria, thus indicating that further iteration to the item pool is required.

Discussion

The current study aimed to develop a comprehensive measure of hallucinatory experiences, across seven modalities, and report preliminary reliability of the scale. Moreover, the authors wished to address a key gap in the trauma-psychosis literature by creating the first self-report measure to explore the phenomenological commonality between hallucinatory experiences and trauma reliving symptoms.

Key findings and suggestions for further QPE development

The QPE gathered data on a wide range of hallucinatory phenomena. Auditory hallucinations were most prominent in the sample, as is echoed in the psychosis literature (Dudley et al., 2023). High rates of visual hallucinations, as well as sensed presence, were also observed; the latter is not commonly included in current measures (Rossell et al., 2019b), thus demonstrating the importance of capturing such phenomena in the QHE. Hallucinations occurred for long durations, were viewed as largely uncontrollable, and caused high levels of distress to participants. Consistent with other reports (Montagnese et al., 2021), multi-modal hallucinations were reported to have more adverse impact on individuals than unimodal ones, negatively affecting functioning, work, and relationships. Most participants (80%) agreed that hallucinatory experiences, across sensory modalities, were linked to traumatic events. Relational links were reported between anomalous experiences and trauma in all domains bar visual hallucinations. This replicates findings from a recent study which demonstrated that relational links, such as experiencing blaming, critical voices associated with past abuse, were most common in a sample of voice-hearers (van den Berg et al., 2023). Identity and content links were also observed, and participants reported the same emotional reaction to hallucinations, and past traumatic events, again highlighting a phenomenological overlap.

Participants' perception that hallucinatory experiences were associated with past traumatic events support findings from other recent studies that have found high levels of content, relational, and identity links between abuse and voice-hearing (Corstens & Longden, 2013; Hardy et al., 2016; Peach et al., 2019; van den Berg et al., 2023). Despite this, this study

found only a weak correlation between the auditory and corresponding trauma subscales. A strong, positive correlation was, however, observed between the visual phenomenology and traumatic experience subscales. Such findings replicate prior research that has reported an association between Post-Traumatic Stress Disorder (PTSD) severity and visual hallucination severity. Significantly higher rates of trauma-related visual hallucinations were reported in the PTSD group, suggesting content that is consistent with the experience of flashbacks(Wearne et al., 2022). Likewise, the strong, positive correlation found between the tactile and corresponding trauma subscales also suggests that participants perceive hallucinations in this domain as resembling trauma re-experiencing symptoms.

Surprisingly, for participants with multi-modal hallucinations, higher severity of hallucinations was associated with lower trauma phenomenology (r=-0.54; 95%CI, -0.85 to -0.03). This finding is in the opposite direction to what is expected given the extant literature on trauma and hallucinations (D'Hondt et al., 2020; Strachan et al., 2023). There is no obvious explanation for this, given the underlying data was checked for outliers and scoring errors. The finding is, however, based on a small sample size (n=19) and there was wide 95% confidence intervals around the estimate. Therefore, caution should be taken in lending any interpretation to this finding without further replication and investigation.

Similarly, preliminary analysis of the QHE's internal consistency demonstrated inconsistent results. Although the measure demonstrated good internal consistency (Cronbach's α >.8), for 6 of the 10 subscales included in this analysis, a further 4 hallucinatory phenomenology subscales (VH, TH, SP and MMH), reported only moderate internal consistency (Cronbach's α ranged from .64-.72). The small sample sizes within each subgroup likely contributed to the wide confidence intervals reported for these subscales, thus limiting confidence in the certainty around these estimates.

Thus, although descriptive statistics indicate a high rate of phenomenological links between hallucinations and trauma, the findings from initial reliability analyses are largely in keeping with the "revise" criteria, indicating that the QHE would benefit from extensive iteration. The authors drew heavily on existing hallucination questionnaires when developing QHE items,

rather than considering theoretical models that might assist conceptualisation of the trauma-psychosis link. Cognitive-Behavioural Models of psychosis propose mechanisms by which trauma may give rise to hallucinations, suggesting that psychosis exists in a continuum of trauma responses(Hardy et al., 2023). For instance, an individual may have a defective schema *"I'm dirty"* represented in their personal semantic memory, because of past psychological abuse. This could result in the generation of a novel, critical voice and accompanied unpleasant olfactory imagery. The resulting distress caused by such experiences could, in turn, disrupt autobiographical memory processing, with the perceptual representation being stored in detail but contextual information inhibited, so that the memory of the voice and smell are later involuntarily re-experienced. Participants' report of content links, in the current study, are in keeping with this hypothesis around the involuntarily retrieval of fragmented, perceptual memories as critical in shaping hallucinations(Hardy, 2017).

Moreover, the dual representation theory suggests that differences between flashbacks and hallucinations may be explained by the severity of dissociative behaviours experiences by an individual (Quidé, 2023). Dissociation during a traumatic event is believed to impede the encoding of memories, leading to the experience of flashbacks(Brewin et al., 1996). However, chronic dissociation, arising to cope with repeated, enduring trauma, may generate more global and complex disturbances in perception, resulting in the manifestation of hallucinations in different modalities(Longden et al., 2012). Correlations between depersonalisation and positive psychotic symptoms further support this proposed pathway underlying trauma and hallucinations (Alderson-Day et al., 2014; Pilton et al., 2015).

Such findings suggest that further iterations of the QHE could include more detailed phenomenological assessment of trauma memory intrusions and anomalous experience intrusions to develop a theoretical understanding of the interplay between specific trauma processes and psychotic symptoms. Indeed, when considering further development of the QHE, the inclusion of items assessing individuals' appraisal of traumatic and hallucinatory experiences, and their physiological-behavioural response, would explore whether a

cognitive-behavioural model of post-traumatic stress in psychosis explains this phenomenological overlap (Hardy, 2017; Morrison et al., 2003; van den Berg et al., 2023).

Strengths and Weaknesses

Few existing measures of hallucinatory experiences include trauma-related items; those that do tend to employ one binary “yes/no” question to ask about the association between hallucinations and unpleasant experiences (Gauntlett-Gilbert & Kuipers, 2005; Rossell et al., 2019a). This study reports the development of the first measure aimed to investigate the phenomenological link between trauma reliving symptoms and hallucinations, thus providing a key addition to the trauma-hallucination literature. Moreover, this measure’s inclusion of a broad range of hallucinatory modalities, (e.g., olfactory, gustatory, multi-modal hallucinations and sensed presence), often overlooked in existing instruments (Rossell et al., 2019a) facilitates a more comprehensive assessment of such experiences across different populations and conditions.

Aside from these strengths, limitations of the study should be considered. The small sample size overall, and especially amongst gustatory and olfactory hallucinatory modality subgroups, limited the statistical analyses that could be performed into the measure’s reliability; all results should therefore be interpreted with caution. Moreover, although individuals who reported the experience of hallucinations in the last year were eligible to participate, the questionnaire was word in such a way that only those who had experienced hallucinations in the last month could be included. This oversight should be addressed in future iterations. The lack of cultural diversity found in the sample is also concerning, especially given the well-established finding that ethnic minority groups are at increased risk of developing psychotic disorders (Jongsma et al., 2021) and often experience more severe PTSD symptoms (Mekawi et al., 2021). Purposive sampling is therefore recommended in further QHE development studies to recruit individuals from a range of backgrounds. Likewise, although the current study consulted key stakeholders in the development of QHE items, only two of these individuals reported lived experience of hallucinations. COnsensus-based Standards for the selection of health Measurement

INstruments (COSMIN) methodology provides guidance for evaluating the content validity of measures, and their development (Terwee et al., 2018). Initial qualitative studies should be conducted with a representative sample of the target population for which the measure is created, to generate relevant items. A cognitive interview study or pilot test with the target population is also recommended to assess the measure's comprehensibility and comprehensiveness. Separate qualitative research with both service users and professionals (e.g., focus groups or individual interviews) is also necessary to collect data to support a measure's content validity. Such studies should again include thorough exploration of item relevance, clarity, and inclusion of all key concepts (Terwee et al., 2018). Thus, before further iterations of the QHE are developed, it is imperative that qualitative research, such as a content validity study, is conducted with a purposive sample of individuals with lived experience of hallucinations to ensure that the measure is an adequate reflection of such phenomena, and acceptable to the target population.

Research and Clinical Implications

The development of the QHE, aimed to facilitate understanding of the trauma-hallucination links, has important implications for future research and clinical practice. Theoretical models which suggest that the same psychological mechanisms (e.g. trauma-related beliefs, episodic memories, cognitive, behavioural and interpersonal emotional regulation) interact to cause PTSD and psychosis also emphasise how these can be viewed as underlying insecure attachment styles (i.e. negative views of self and others, with emotional regulation problems), which are also characteristic of Complex-Post Traumatic Stress Disorder (CPTSD Hardy et al., 2023). Indeed, a current clinical trial aims to investigate if CPTSD is casually related to psychosis using psychological interventions aimed to treat different symptoms of CPTSD (emotion dysregulation, negative self-concept, and interpersonal difficulties) (NCT05281640). Adaptation of the QHE to include items specific to CPTSD symptoms would aid exploration of how these affect hallucinations, thus improving the conceptualisation of psychosis. Currently, research into the development of psychological treatment for psychosis has focused on evaluating different trauma interventions with this

population (Peters et al., 2022; Varese et al., 2021). It is anticipated that, following further iteration, the QHE will facilitate increased theoretical understanding into the interplay between trauma-related experiences and symptoms those found in psychosis. This will assist researchers in the evaluation of more targeted treatment for specific maintenance factors, such as dissociation, intrusions, and negative schema. For the clinician, we hope the measure will begin a conversation with individuals who report hallucinations, about experiences of trauma, and so support them to make sense of their difficulties and consider options for intervention.

References

Airey, N. D., Taylor, C. D. J., Vikram, A., & Berry, K. (2023). Trauma measures for use with psychosis populations: A systematic review of psychometric properties using COSMIN. *Psychiatry Res*, *323*, 115163. https://doi.org/10.1016/j.psychres.2023.115163

Alameda, L., Christy, A., Rodriguez, V., Salazar de Pablo, G., Thrush, M., Shen, Y., Alameda, B., Spinazzola, E., Iacoponi, E., Trotta, G., Carr, E., Ruiz Veguilla, M., Aas, M., Morgan, C., & Murray, R. M. (2021). Association Between Specific Childhood Adversities and Symptom Dimensions in People With Psychosis: Systematic Review and Meta-Analysis. *Schizophr Bull*, *47*(4), 975-985. https://doi.org/10.1093/schbul/sbaa199

Alderson-Day, B., McCarthy-Jones, S., Bedford, S., Collins, H., Dunne, H., Rooke, C., & Fernyhough, C. (2014). Shot through with voices: Dissociation mediates the relationship between varieties of inner speech and auditory hallucination proneness. *Consciousness and Cognition*, *27*, 288-296. https://doi.org/https://doi.org/10.1016/j.concog.2014.05.010

Allen, S., Goodall, T., Jones, C., James, R., & Surtees, A. (2023). What Is the Prevalence of Visual Hallucinations in a First-Episode Psychosis Population? A Systematic Review and Meta-analysis of the Literature. *Schizophrenia Bulletin Open*, *4*(1). https://doi.org/10.1093/schizbullopen/sgad002

Andreasen, N. C. (1984). Scale for the assessment of positive symptoms. *Psychiatrie & Psychobiologie*.

Aynsworth, C., Collerton, D., & Dudley, R. (2017). Measures of visual hallucinations: Review and recommendations [Empirical Study; Literature Review; Qualitative Study]. *Clinical Psychology Review*, *57*, 164-182. https://doi.org/https://dx.doi.org/10.1016/j.cpr.2017.05.001

Aynsworth, C., Rolinson, J., Pervez, M., Collerton, D., & Dudley, R. (2022). What is the frequency and nature of visual hallucinations in non-clinical participants? *Psychology and psychotherapy.*, *11*. https://doi.org/https://dx.doi.org/10.1111/papt.12440

Aynsworth, C., Rolinson, J., Pervez, M., Collerton, D., & Dudley, R. (2023). What is the frequency and nature of visual hallucinations in non-clinical participants? [Research Support, Non-U.S. Gov't]. *Psychology & Psychotherapy: Theory, Research & Practice*, *96*(2), 281-295. https://doi.org/https://dx.doi.org/10.1111/papt.12440

Bailey, T., Alvarez-Jimenez, M., Garcia-Sanchez, A. M., Hulbert, C., Barlow, E., & Bendall, S. (2018). Childhood Trauma Is Associated With Severity of Hallucinations and Delusions in Psychotic Disorders: A Systematic Review and Meta-Analysis. *Schizophr Bull*, *44*(5), 1111-1122. https://doi.org/10.1093/schbul/sbx161

Beards, S., Gayer-Anderson, C., Borges, S., Dewey, M. E., Fisher, H. L., & Morgan, C. (2013). Life events and psychosis: a review and meta-analysis. *Schizophr Bull*, *39*(4), 740-747. https://doi.org/10.1093/schbul/sbt065

Bell, V., Halligan, P. W., & Ellis, H. D. (2006). The Cardiff Anomalous Perceptions Scale (CAPS): a new validated measure of anomalous perceptual experience. *Schizophrenia Bulletin*, *32*(2), 366-377.

Bell, V., Raballo, A., Larøi, F., & Aleman, A. (2010). Assessment of hallucinations. *Hallucinations: A guide to treatment and management*, 377-397.

Bendall, S., Jackson, H. J., Hulbert, C. A., & McGorry, P. D. (2008). Childhood trauma and psychotic disorders: a systematic, critical review of the evidence. *Schizophr Bull*, *34*(3), 568-579. https://doi.org/10.1093/schbul/sbm121

Brewin, C. R., Dalgleish, T., & Joseph, S. (1996). A dual representation theory of posttraumatic stress disorder. *Psychol Rev*, *103*(4), 670-686. https://doi.org/10.1037/0033-295x.103.4.670

Černis, E., Cooper, M., & Chan, C. (2018). Developing a new measure of dissociation: The Dissociative Experiences Measure, Oxford (DEMO). *Psychiatry Research*, *269*, 229-236. https://doi.org/https://doi.org/10.1016/j.psychres.2018.08.060

Chouinard, G., & Miller, R. (1999). A rating scale for psychotic symptoms (RSPS): Part I: theoretical principles and subscale 1: perception symptoms (illusions and hallucinations). *Schizophrenia Research*, *38*(2-3), 101-122.

Chouinard, V.-A., Shinn, A. K., Valeri, L., Chouinard, P. A., Gardner, M. E., Asan, A., Cohen, B. M., & Ongur, D. (2019). Visual hallucinations associated with multimodal hallucinations, suicide attempts and morbidity of illness in psychotic disorders [Empirical Study; Quantitative Study]. *Schizophrenia Research*, *208*, 196-201. https://doi.org/https://dx.doi.org/10.1016/j.schres.2019.02.022

Clark, L. A., & Watson, D. (2016). *Constructing validity: Basic issues in objective scale development* [doi:10.1037/14805-012]. American Psychological Association. https://doi.org/10.1037/14805-012

Clark, L. A., & Watson, D. (2019). Constructing validity: New developments in creating objective measuring instruments. *Psychol Assess*, *31*(12), 1412-1427. https://doi.org/10.1037/pas0000626

Clark, M. L., Waters, F., Vatskalis, T. M., & Jablensky, A. (2017). On the interconnectedness and prognostic value of visual and auditory hallucinations in first-episode psychosis. *European Psychiatry*, *41*, 122-128. https://doi.org/https://doi.org/10.1016/j.eurpsy.2016.10.011

Corstens, D., & Longden, E. (2013). The origins of voices: links between life history and voice hearing in a survey of 100 cases. *Psychosis*, *5*(3), 270-285. https://doi.org/10.1080/17522439.2013.816337

D'Hondt, F., Jardri, R., Medjkane, F., Notredame, C.-E., Sharkey, L., & Vaiva, G. (2020). Association between childhood trauma and multimodal early-onset hallucinations. *The British Journal of Psychiatry*, *216*(3), 156-158. https://doi.org/10.1192/bjp.2019.266

de Chazeron, I., Pereira, B., Chereau-Boudet, I., Brousse, G., Misdrahi, D., Fenelon, G., Tronche, A. M., Schwan, R., Lancon, C., Marques, A., Debilly, B., Durif, F., & Llorca, P. M. (2015). Validation of a Psycho-Sensory hAllucinations Scale (PSAS) in schizophrenia and Parkinson's disease. *Schizophrenia Research*, *161*(2), 269-276. https://doi.org/doi:https://dx.doi.org/10.1016/j.schres.2014.11.010

DeVellis, R. F. (2016). *Scale Development: Theory and Applications*. SAGE Publications. https://books.google.co.uk/books?id=48ACCwAAQBAJ

Dudley, R., Collerton, D., Nicholson, M., & Mosimann, U. (2013). Clinical characteristics of disclosed visual hallucinations in users of an Early Intervention in Psychosis

Service [Empirical Study; Interview; Quantitative Study]. *Psychosis: Psychological, Social and Integrative Approaches*, *5*(2), 127-133. https://doi.org/https://dx.doi.org/10.1080/17522439.2012.699543

Dudley, R., Watson, F., O'Grady, L., Aynsworth, C., Dodgson, G., Common, S., Day, B. A., & Fernyhough, C. (2023). Prevalence and nature of multi-sensory and multi-modal hallucinations in people with first episode psychosis [Research Support, Non-U.S. Gov't]. *Psychiatry Research*, *319*, 114988. https://doi.org/https://dx.doi.org/10.1016/j.psychres.2022.114988

Franceschi, P. (2011). On the plausibility of psychotic hallucinations. *Activitas Nervosa Superior*, *53*, 35-43. https://doi.org/doi:http://dx.doi.org/10.1007/BF03379932

Gauntlett-Gilbert, J., & Kuipers, E. (2005). Visual hallucinations in psychiatric conditions: appraisals and their relationship to distress. *British Journal of Clinical Psychology*, *44*(1), 77-87. https://doi.org/doi:10.1348/014466504x19451

Ghanem, M., Evangeli-Dawson, C., & Georgiades, A. (2023). The role of culture on the phenomenology of hallucinations and delusions, explanatory models, and help-seeking attitudes: A narrative review. *Early Intervention in Psychiatry*, *17*(9), 843-863. https://doi.org/https://doi.org/10.1111/eip.13449

Goetz, C. G., Tilley, B. C., Shaftman, S. R., Stebbins, G. T., Fahn, S., Martinez-Martin, P., Poewe, W., Sampaio, C., Stern, M. B., Dodel, R., Dubois, B., Holloway, R., Jankovic, J., Kulisevsky, J., Lang, A. E., Lees, A., Leurgans, S., LeWitt, P. A., Nyenhuis, D., . . . LaPelle, N. (2008). Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Mov Disord*, *23*(15), 2129-2170. https://doi.org/10.1002/mds.22340

Haddock, G., McCarron, J., Tarrier, N., & Faragher, E. B. (1999a). Scales to measure dimensions of hallucinations and delusions: the psychotic symptom rating scales (PSYRATS). *Psychological Medicine*, *29*(4), 879-889. https://ovidsp.ovid.com/ovidweb.cgi?T=JS&CSC=Y&NEWS=N&PAGE=fulltext&D=med4&AN=10473315

Haddock, G., McCarron, J., Tarrier, N., & Faragher, E. B. (1999b). Scales to measure dimensions of hallucinations and delusions: the psychotic symptom rating scales (PSYRATS). *Psychological Medicine*, *29*(4), 879-889. https://doi.org/10.1017/S0033291799008661

Hardy, A. (2017). Pathways from Trauma to Psychotic Experiences: A Theoretically Informed Model of Posttraumatic Stress in Psychosis [Hypothesis and Theory]. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00697

Hardy, A., Emsley, R., Freeman, D., Bebbington, P., Garety, P. A., Kuipers, E. E., Dunn, G., & Fowler, D. (2016). Psychological Mechanisms Mediating Effects Between Trauma and Psychotic Symptoms: The Role of Affect Regulation, Intrusive Trauma Memory, Beliefs, and Depression. *Schizophrenia Bulletin*, *42*(suppl_1), S34-S43. https://doi.org/10.1093/schbul/sbv175

Hardy, A., Keen, N., van den Berg, D., Varese, F., Longden, E., Ward, T., & Brand, R. M. (2023). Trauma therapies for psychosis: A state-of-the-art review. *Psychol Psychother*. https://doi.org/10.1111/papt.12499

Heriot-Maitland, C., Vitoratou, S., Peters, E., Hermans, K., Wykes, T., & Brett, C. (2023). Detecting anomalous experiences in the community: The Transpersonal

Experiences Questionnaire (TEQ) . *Psychology and Psychotherapy: Theory, Research and Practice*, *96*(2), 383-398. https://doi.org/10.1111/papt.12445

Johns, L. C., & van Os, J. (2001). The continuity of psychotic experiences in the general population. *Clinical Psychology Review*, *21*(8), 1125-1141. https://doi.org/10.1016/s0272-7358(01)00103-9

Jongsma, Gayer-Anderson, C., Tarricone, I., Velthorst, E., van der Ven, E., Quattrone, D., di Forti, M., Menezes, P. R., Del-Ben, C. M., Arango, C., Lasalvia, A., Berardi, D., La Cascia, C., Bobes, J., Bernardo, M., Sanjuán, J., Santos, J. L., Arrojo, M., de Haan, L., . . . Kirkbride, J. B. (2021). Social disadvantage, linguistic distance, ethnic minority status and first-episode psychosis: results from the EU-GEI case–control study. *Psychological Medicine*, *51*(9), 1536-1548. https://doi.org/10.1017/S003329172000029X

Jongsma, H. E., Karlsen, S., Kirkbride, J. B., & Jones, P. B. (2021). Understanding the excess psychosis risk in ethnic minorities: the impact of structure and identity. *Soc Psychiatry Psychiatr Epidemiol*, *56*(11), 1913-1921. https://doi.org/10.1007/s00127-021-02042-8

Justo-Núñez, M., Morris, L., & Berry, K. (2022). Self-report measures of secure attachment in adulthood: A systematic review. *Clin Psychol Psychother*, *29*(6), 1812-1842. https://doi.org/10.1002/cpp.2756

Kay, Fiszbein, A., & Opler, L. A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin*, *13*(2), 261-276.

Kraan, T., Velthorst, E., Smit, F., de Haan, L., & van der Gaag, M. (2015). Trauma and recent life events in individuals at ultra high risk for psychosis: Review and meta-analysis. *Schizophrenia Research*, *161*(2), 143-149. https://doi.org/https://doi.org/10.1016/j.schres.2014.11.026

Lai, S., Bruce, V., & Collerton, D. (2016). Visual Hallucinations in Older People: Appraisals but not Content or Phenomenology Influence Distress. *Behavioural and Cognitive Psychotherapy*, *44*, 705-710.

Laloyaux, J., Bless, J. J., Hugdahl, K., Kråkvik, B., Vedul-Kjelsås, E., Kalhovde, A. M., & Larøi, F. (2019). Multimodal hallucinations are associated with poor mental health and negatively impact auditory hallucinations in the general population: Results from an epidemiological study. *Schizophr Res*, *210*, 319-322. https://doi.org/10.1016/j.schres.2019.06.005

Larøi, F., Marczewski, P., & Van der Linden, M. (2004). Further evidence of the multi-dimensionality of hallucinatory predisposition: factor structure of a modified version of the Launay-Slade Hallucinations Scale in a normal sample. *European Psychiatry: the Journal of the Association of European Psychiatrists*, *19*(1), 15-20. https://doi.org/10.1016/s0924-9338(03)00028-2

Larøi, F., & Van der Linden, M. (2005). Metacognitions in proneness towards hallucinations and delusions. *Behaviour Research & Therapy*, *43*(11), 1425-1441. https://doi.org/10.1016/j.brat.2004.10.008

Launay, G., & Slade, P. (1981). The measurement of hallucinatory predisposition in male and female prisoners. *Personality and Individual Differences*, *2*(3), 221-234. https://doi.org/https://doi.org/10.1016/0191-8869(81)90027-1

Linszen, M. M. J., de Boer, J. N., Schutte, M. J. L., Begemann, M. J. H., de Vries, J., Koops, S., Blom, R. E., Bohlken, M. M., Heringa, S. M., Blom, J. D., & Sommer, I. E. C. (2022). Occurrence and phenomenology of hallucinations in the general

population: A large online survey. *Schizophrenia*, *8*(1), 41. https://doi.org/10.1038/s41537-022-00229-9

Longden, E., Madill, A., & Waterman, M. G. (2012). Dissociation, trauma, and the role of lived experience: toward a new conceptualization of voice hearing [Review]. *Psychological bulletin*, *138*(1), 28-76. https://doi.org/https://dx.doi.org/10.1037/a0025995

Mason, O., Claridge, G., & Jackson, M. (1995). New scales for the assessment of schizotypy. *Personality and Individual Differences*, *18*(1), 7-13.

Mason, O., Linney, Y., & Claridge, G. (2005). Short scales for measuring schizotypy. *Schizophrenia Research*, *78*(2), 293-296. https://doi.org/https://doi.org/10.1016/j.schres.2005.06.020

McCarthy-Jones, S., Smailes, D., Corvin, A., Gill, M., Morris, D. W., Dinan, T. G., Murphy, K. C., Anthony O'Neill, F., Waddington, J. L., Australian Schizophrenia Research, B., Donohoe, G., & Dudley, R. (2017). Occurrence and co-occurrence of hallucinations by modality in schizophrenia-spectrum disorders. *Psychiatry Research*, *252*, 154-160. https://doi.org/https://doi.org/10.1016/j.psychres.2017.01.102

McCarthy-Jones, S., Thomas, N., Strauss, C., Dodgson, G., Jones, N., Woods, A., Brewin, C. R., Hayward, M., Stephane, M., Barton, J., Kingdon, D., & Sommer, I. E. (2014). Better than mermaids and stray dogs? Subtyping auditory verbal hallucinations and its implications for research and practice. *Schizophr Bull*, *40 Suppl 4*(Suppl 4), S275-284. https://doi.org/10.1093/schbul/sbu018

Meadows, K. (2021). Cognitive Interviewing Methodologies. *Clinical Nursing Research*, *30*(4), 375-379. https://doi.org/10.1177/10547738211014099

Mekawi, Y., Carter, S., Brown, B., Martinez de Andino, A., Fani, N., Michopoulos, V., & Powers, A. (2021). Interpersonal Trauma and Posttraumatic Stress Disorder among Black Women: Does Racial Discrimination Matter? *Journal of Trauma & Dissociation*, *22*(2), 154-169. https://doi.org/10.1080/15299732.2020.1869098

Mitchell, C. A. A., Maybery, M. T., Russell-Smith, S. N., Collerton, D., Gignac, G. E., & Waters, F. (2017). The Structure and Measurement of Unusual Sensory Experiences in Different Modalities: The Multi-Modality Unusual Sensory Experiences Questionnaire (MUSEQ) [Original Research]. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.01363

Mokkink, L. B., De Vet, H. C., Prinsen, C. A., Patrick, D. L., Alonso, J., Bouter, L. M., & Terwee, C. B. (2018). COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Quality of Life Research*, *27*, 1171-1179.

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & De Vet, H. C. (2010a). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of Life Research*, *19*, 539-549.

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. (2010b). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of clinical epidemiology*, *63*(7), 737-745.

Montagnese, M., Leptourgos, P., Fernyhough, C., Waters, F., Larøi, F., Jardri, R., McCarthy-Jones, S., Thomas, N., Dudley, R., Taylor, J.-P., Collerton, D., &

Urwyler, P. (2020). A Review of Multimodal Hallucinations: Categorization, Assessment, Theoretical Perspectives, and Clinical Recommendations. *Schizophrenia Bulletin*, *47*(1), 237-248. https://doi.org/10.1093/schbul/sbaa101

Montagnese, M., Leptourgos, P., Fernyhough, C., Waters, F., Larøi, F., Jardri, R., McCarthy-Jones, S., Thomas, N., Dudley, R., Taylor, J. P., Collerton, D., & Urwyler, P. (2021). A Review of Multimodal Hallucinations: Categorization, Assessment, Theoretical Perspectives, and Clinical Recommendations. *Schizophr Bull*, *47*(1), 237-248. https://doi.org/10.1093/schbul/sbaa101

Morrison, A. P., Frame, L., & Larkin, W. (2003). Relationships between trauma and psychosis: a review and integration. *Br J Clin Psychol*, *42*(Pt 4), 331-353. https://doi.org/10.1348/014466503322528892

Morrison, A. P., Wells, A., & Nothard, S. (2000). Cognitive factors in predisposition to auditory and visual hallucinations. *Br J Clin Psychol*, *39*(1), 67-78. https://doi.org/10.1348/014466500163112

Mosimann, U. P., Collerton, D., Dudley, R., Meyer, T. D., Graham, G., Dean, J. L., Bearn, D., Killen, A., Dickinson, L., Clarke, M. P., McKeith, I. G., Mosimann, U. P., Collerton, D., Dudley, R., Meyer, T. D., Graham, G., Dean, J. L., Bearn, D., Killen, A., & Dickinson, L. (2008). A semi-structured interview to assess visual hallucinations in older people. *International Journal of Geriatric Psychiatry*, *23*(7), 712-718. https://doi.org/10.1002/gps.1965

Muller, A. J., Mills, J. M. Z., O'Callaghan, C., Naismith, S. L., Clouston, P. D., Lewis, S. J. G., & Shine, J. M. (2018). Informant- and Self-Appraisals on the Psychosis and Hallucinations Questionnaire (PsycH-Q) Enhances Detection of Visual Hallucinations in Parkinson's Disease. *Movement Disorders Clinical Practice*, *5*(6), 607-613. https://doi.org/https://dx.doi.org/10.1002/mdc3.12683

Oorschot, M., Lataster, T., Thewissen, V., Bentall, R., Delespaul, P., & Myin-Germeys, I. (2012). Temporal dynamics of visual and auditory hallucinations in psychosis. *Schizophrenia Research*, *140(1-3)*, 77-82. https://doi.org/https://dx.doi.org/10.1016/j.schres.2012.06.010

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., . . . Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, *372*, n71. https://doi.org/10.1136/bmj.n71

Peach, N., Alvarez-Jimenez, M., Cropper, S. J., Sun, P., & Bendall, S. (2019). Testing models of post-traumatic intrusions, trauma-related beliefs, hallucinations, and delusions in a first episode psychosis sample. *British Journal of Clinical Psychology*, *58*(2), 154-172. https://doi.org/https://doi.org/10.1111/bjc.12206

Peters, E., Hardy, A., Dudley, R., Varese, F., Greenwood, K., Steel, C., Emsley, R., Keen, N., Bowe, S., Swan, S., Underwood, R., Longden, E., Byford, S., Potts, L., Heslin, M., Grey, N., Turkington, D., Fowler, D., Kuipers, E., & Morrison, A. (2022). Multisite randomised controlled trial of trauma-focused cognitive behaviour therapy for psychosis to reduce post-traumatic stress symptoms in people with co-morbid post-traumatic stress disorder and psychosis, compared to treatment as usual: study protocol for the STAR (Study of Trauma And Recovery) trial. *Trials [Electronic Resource]*, *23*(1), 429. https://doi.org/10.1186/s13063-022-06215-x

Pilton, M., Varese, F., Berry, K., & Bucci, S. (2015). The relationship between dissociation and voices: A systematic literature review and meta-analysis. *Clinical Psychology Review*, *40*, 138-155. https://doi.org/https://doi.org/10.1016/j.cpr.2015.06.004

Prinsen, Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., de Vet, H. C. W., & Terwee, C. B. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research*, *27*(5), 1147-1157. https://doi.org/10.1007/s11136-018-1798-3

Prinsen, C. A., Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., De Vet, H. C., & Terwee, C. B. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research*, *27*, 1147-1157.

Quidé, Y. (2023). Dissociation, trauma and the experience of visual hallucinations in post-traumatic stress disorder and schizophrenia. *BJPsych Open*, *9*(1), e23. https://doi.org/10.1192/bjo.2023.3

Read, J., van Os, J., Morrison, A. P., & Ross, C. A. (2005). Childhood trauma, psychosis and schizophrenia: a literature review with theoretical and clinical implications. *Acta Psychiatr Scand*, *112*(5), 330-350. https://doi.org/10.1111/j.1600-0447.2005.00634.x

Rossell, S. L., Schutte, M. J. L., Toh, W. L., Thomas, N., Strauss, C., Linszen, M. M. J., van Dellen, E., Heringa, S. M., Teunisse, R., Slotema, C. W., & Sommer, I. E. C. (2019a). The Questionnaire for Psychotic Experiences: An Examination of the Validity and Reliability. *Schizophrenia Bulletin*, *45*(45), S78-S87. https://doi.org/doi:https://dx.doi.org/10.1093/schbul/sby148

Rossell, S. L., Schutte, M. J. L., Toh, W. L., Thomas, N., Strauss, C., Linszen, M. M. J., van Dellen, E., Heringa, S. M., Teunisse, R., Slotema, C. W., & Sommer, I. E. C. (2019b). The Questionnaire for Psychotic Experiences: An Examination of the Validity and Reliability [Research Support, Non-U.S. Gov't]. *Schizophrenia Bulletin*, *45*(45 Suppl 1), S78-S87. https://doi.org/https://dx.doi.org/10.1093/schbul/sby148

Ryan, K., Gannon-Slater, N., & Culbertson, M. J. (2012). Improving Survey Methods With Cognitive Interviews in Small- and Medium-Scale Evaluations. *American Journal of Evaluation*, *33*(3), 414-430. https://doi.org/10.1177/1098214012441499

Schutte, M. J., Linszen, M. M., Marschall, T. M., ffytche, D. H., Koops, S., van Dellen, E., Heringa, S. M., Slooter, A. J., Teunisse, R., van den Heuvel, O. A., Lemstra, A. W., Foncke, E. M., Slotema, C. W., de Jong, J., Rossell, S. L., & Sommer, I. E. (2020). Hallucinations and other psychotic experiences across diagnoses: A comparison of phenomenological features [Empirical Study; Quantitative Study]. *Psychiatry Research Vol 292 2020, ArtID 113314*, *292*. https://doi.org/https://dx.doi.org/10.1016/j.psychres.2020.113314

Seiler, N., Nguyen, T., Yung, A., & O'Donoghue, B. (2020). Terminology and assessment tools of psychosis: A systematic narrative review. *Psychiatry & Clinical Neurosciences*, *74*(4), 226-246. https://doi.org/10.1111/pcn.12966

Shine, J. M., Mills, J. M. Z., Qiu, J., O'Callaghan, C., Terpening, Z., Halliday, G. M., Naismith, S. L., & Lewis, S. J. G. (2015). Validation of the Psychosis and Hallucinations Questionnaire in Non-demented Patients with Parkinson's Disease. *Movement Disorders Clinical Practice*, *2(2)*, 175-181. https://doi.org/https://dx.doi.org/10.1002/mdc3.12139

Smith, E. L., Garety, P. A., Harding, H., & Hardy, A. (2021). Are there reliable and valid measures of anxiety for people with psychosis? A systematic review of psychometric properties. *Psychol Psychother*, *94*(1), 173-198. https://doi.org/10.1111/papt.12265

Strachan, L. P., Paulik, G., Preece, D. A., & McEvoy, P. M. (2023). Pathways from trauma to unusual perceptual experiences: Modelling the roles of insecure attachment,

negative affect, emotion regulation and dissociation. *Psychology and Psychotherapy: Theory, Research and Practice*, *96*(4), 934-951. https://doi.org/https://doi.org/10.1111/papt.12486

Subhi, Y., Schmidt, D. C., Bach-Holm, D., Kolko, M., & Singh, A. (2021). Prevalence of Charles Bonnet syndrome in patients with glaucoma: a systematic review with meta-analyses [Review]. *Acta Ophthalmologica*, *99(2)*, 128-133. https://doi.org/https://dx.doi.org/10.1111/aos.14567

Taitimu, M., Read, J., & McIntosh, T. (2018). Ngā Whakāwhitinga (standing at the crossroads): How Māori understand what Western psychiatry calls "schizophrenia". *Transcultural Psychiatry*, *55*(2), 153-177. https://doi.org/10.1177/1363461518757800

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *Int J Med Educ*, *2*, 53-55. https://doi.org/10.5116/ijme.4dfb.8dfd

Terwee, C. B., Bot, S. D., de Boer, M. R., van der Windt, D. A., Knol, D. L., Dekker, J., Bouter, L. M., & de Vet, H. C. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of clinical epidemiology*, *60*(1), 34-42.

Terwee, C. B., Bot, S. D., de Boer, M. R., van der Windt, D. A., Knol, D. L., Dekker, J., Bouter, L. M., & de Vet, H. C. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*, *60*(1), 34-42. https://doi.org/10.1016/j.jclinepi.2006.03.012

Terwee, C. B., Mokkink, L. B., Knol, D. L., Ostelo, R. W., Bouter, L. M., & de Vet, H. C. (2012). Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Quality of Life Research*, *21*, 651-657.

Terwee, C. B., Prinsen, C. A. C., Chiarotto, A., Westerman, M. J., Patrick, D. L., Alonso, J., Bouter, L. M., de Vet, H. C. W., & Mokkink, L. B. (2018). COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Quality of Life Research*, *27*(5), 1159-1170. https://doi.org/10.1007/s11136-018-1829-0

Toh, W. L., Moseley, P., & Fernyhough, C. (2022). Hearing voices as a feature of typical and psychopathological experience. *Nature Reviews Psychology*, *1*(2), 72-86. https://doi.org/10.1038/s44159-021-00013-z

Urwyler, P., Nef, T., Killen, A., Collerton, D., Thomas, A., Burn, D., McKeith, I., & Mosimann, U. P. (2014). Visual complaints and visual hallucinations in Parkinson's disease. *Parkinsonism and Related Disorders*, *20(3)*, 318-322. https://doi.org/https://dx.doi.org/10.1016/j.parkreldis.2013.12.009

van den Berg, D., Tolmeijer, E., Jongeneel, A., Staring, A. B. P., Palstra, E., van der Gaag, M., & Hardy, A. (2023). Voice phenomenology as a mirror of the past. *Psychol Med*, *53*(7), 2954-2962. https://doi.org/10.1017/s0033291721004955

van Ommen, M., van Laar, T., Cornelissen, F., & Bruggeman, R. (2019). Visual hallucinations in psychosis. *Psychiatry Research Vol 280 2019, ArtID 112517*, *280*. https://doi.org/doi:https://dx.doi.org/10.1016/j.psychres.2019.112517

Varese, F., Sellwood, W., Aseem, S., Awenat, Y., Bird, L., Bhutani, G., Carter, L.-A., Davies, L., Davis, C., Horne, G., Keane, D., Logie, R., Malkin, D., Potter, F., van den Berg, D., Zia, S., & Bentall, R. (2021). Eye movement desensitization and reprocessing therapy for psychosis (EMDRp): Protocol of a feasibility randomized

controlled trial with early intervention service users. *Early Intervention in Psychiatry*, *15*(5), 1224-1233. https://doi.org/https://doi.org/10.1111/eip.13071

Varese, F., Smeets, F., Drukker, M., Lieverse, R., Lataster, T., Viechtbauer, W., Read, J., Van Os, J., & Bentall, R. P. (2012). Childhood adversities increase the risk of psychosis: a meta-analysis of patient-control, prospective-and cross-sectional cohort studies. *Schizophrenia Bulletin*, *38*(4), 661-671.

Waters, F., Blom, J. D., Jardri, R., Hugdahl, K., & Sommer, I. E. C. (2018). Auditory hallucinations, not necessarily a hallmark of psychotic disorder. *Psychol Med*, *48*(4), 529-536. https://doi.org/10.1017/s0033291717002203

Waters, F., Collerton, D., Ffytche, D. H., Jardri, R., Pins, D., Dudley, R., Blom, J. D., Mosimann, U. P., Eperjesi, F., Ford, S., & Laroi, F. (2014). Visual hallucinations in the psychosis spectrum and comparative information from neurodegenerative disorders and eye disease. *Schizophrenia Bulletin*, *40(SUPPL. 4)*, S233-S245. https://doi.org/https://dx.doi.org/10.1093/schbul/sbu036

Wearne, D., Ayalde, J., Curtis, G., Gopisetty, A., Banerjee, A., Melvill-Smith, P., Orr, K., Rajanthiran, L., & Waters, F. (2022). Visual phenomenology in schizophrenia and post-traumatic stress disorder: an exploratory study. *BJPsych Open*, *8*(4), e143. https://doi.org/10.1192/bjo.2022.544

Wilson, R., Collerton, D., Freeston, M., Christodoulides, T., & Dudley, R. (2016). Is Seeing Believing? The Process of Change During Cognitive-behavioural Therapy for Distressing Visual Hallucinations [Case Reports]. *Clinical Psychology & Psychotherapy*, *23*(4), 285-297. https://doi.org/https://dx.doi.org/10.1002/cpp.1962

Wood, A. M., & Boyce, C. J. (2017). Developing, Evaluating, and Using Subjective Scales of Personality, Preferences, and Well-Being: A Guide to Psychometrics for Psychologists and Economists. In *Economic Psychology* (pp. 88-103). https://doi.org/https://doi.org/10.1002/9781118926352.ch6

Yates, K., Lång, U., Peters, E. M., Wigman, J. T. W., Boyda, D., McNicholas, F., Cannon, M., Alderson-Day, B., Bloomfield, M., Ramsay, H., & Kelleher, I. (2022). Hallucinations as a risk marker for suicidal behaviour in individuals with a history of sexual assault: a general population study with instant replication. *Psychol Med*, 1-7. https://doi.org/10.1017/s0033291722001532

Appendix 1.1: PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	7
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	8
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	10
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	11
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	13-14
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	12
Search strategy	7	Present the full search strategies for all databases, registers, and websites, including any filters and limits used.	Appendix 1.2
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	13-14, 19
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	13-14,20

Section and Topic	Item #	Checklist item	Location where item is reported
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g., for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	N/A
	10b	List and define all other variables for which data were sought (e.g., participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	14, 21-24
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	15-17
Effect measures	12	Specify for each outcome the effect measure(s) (e.g., risk ratio, mean difference) used in the synthesis or presentation of results.	N/A
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g., tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	N/A
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	N/A
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	N/A
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	N/A
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g., subgroup analysis, meta-regression).	N/A
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	N/A

Section and Topic	Item #	Checklist item	Location where item is reported
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	N/A
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	N/A
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	Figure 1.1
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	N/A
Study characteristics	17	Cite each included study and present its characteristics.	Table 1.2, Appendix 1.3
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	Table 1.3, 1.4, Appendix 1.7, 1,8
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g., confidence/credible interval), ideally using structured tables or plots.	N/A
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	N/A
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g., confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	N/A
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	N/A
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	N/A

Section and Topic	Item #	Checklist item	Location where item is reported
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	N/A
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	Table 1.4, Appendix 1.7-1.9
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	36-37
	23b	Discuss any limitations of the evidence included in the review.	38-39
	23c	Discuss any limitations of the review processes used.	38-39
	23d	Discuss implications of the results for practice, policy, and future research.	39-40
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	12
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	12
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	12
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	N/A
Competing interests	26	Declare any competing interests of review authors.	N/A
Availability of data, code, and other materials	27	Report which of the following are publicly available and where they can be found template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	Appendix 1.8

f

Appendix 1.2: Search Strategy

Ovid (APA PsycInfo 1806-Present).	
S1 Visual Hallucinations	visual hallucinations/
S2 Visual Hallucinations	(visual* hallucinat*). ti,ab.
S3 Measures	(scale* or measure* or interview* or questionnaire* or "self-report" or prevalence or survey* or task or tool* or checklist* or "outcome measure*" or valid* or reliab* or psychometric* or reproducibility or "factor analy*" or "internal consistency"). ti,ab.
S4 Measures	Exp Psychometrics/
S5 Measures	exp Interview Schedules/
S6 Measures	Exp Factor Analysis/
S7	**S1 OR S2**
S8	**S3 OR S4 OR S5 OR S6**
S9	**S7 AND S8**
S10	**Limit S9 to English language**
S11	**Limit S10 to 2016 to present**

OVID (MEDLINE (R) 1946 to Present.	
S1 Visual Hallucinations	(visual* hallucinat*). ti,ab
S2 Measures	Psychometrics/
S3 Measures	exp "Surveys and Questionnaires"

S4 Measures	Factor Analysis, Statistical/
S5 Measures	“Reproducibility of Results”/
S6 Measures	prevalence/
S7 Measures	(scale* or measure* or interview* or questionnaire* or “self-report” or prevalence or survey* or task or tool* or checklist* or “outcome measure*” or valid* or reliab* or psychometric* or reproducibility or “factor analy*” or “internal consistency”). ti, ab
S8	**S2 OR S3 OR S4 OR S5 OR S6 OR S7**
S9	**S1 AND S8**
S10	**Limit S9 to English language**
S11	**Limit S10 to 2016 to present**

OVID (Embase 1947- Present.)	
S1 Visual Hallucinations	visual hallucination/
S2 Visual Hallucinations	(visual* hallucinat*). ti, ab
S3 Measures	questionnaire/
S4 Measures	Exp reliability/
S5 Measures	Exp validity/
S6 Measures	psychometry/
S7 Measures	Exp factor analysis/
S8 Measures	Prevalence/
S9 Measures	(scale* or measure* or interview* or questionnaire* or “self-report” or prevalence or survey* or task or tool* or checklist* or “outcome measure*” or valid* or reliab* or psychometric* or reproducibility or “factor analy*” or “internal consistency”). ti, ab.
S10	**S1 OR S2**
S11	**S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9**
S12	**S10 AND S11**
S13	**Limit S12 to English language**

S14	**Limit S13 to 2016 to present**

EBSCOhost (CINAHL).	
S1 Visual Hallucinations	TI visual* hallucinat* or AB visual* hallucinat*
S2 Measures	(MH “Questionnaires+”)
S3 Measures	(MM “Reliability and Validity”)
S4 Measures	(MM “Psychometrics”)
S5 Measures	(MH “Factor Analysis+”)
S6 Measures	(MM “Prevalence”)
S7 Measures	(MM “Instrument Validation”)
S8 Measures	TI (scale* or measure* or interview* or questionnaire* or “self-report” or prevalence or survey* or task or tool* or checklist* or “outcome measure*” or valid* or reliab* or psychometric* or reproducibility or “factor analy*” or “internal consistency”) or AB (scale* or measure* or interview* or questionnaire* or “self-report” or prevalence or survey* or task or tool* or checklist* or “outcome measure*” or valid* or reliab* or psychometric* or reproducibility or “factor analy*” or “internal consistency”)
S9	**S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8**
S10	**S1 AND S9**
S11	**Limit S10 to English language**
S12	**Limit S11 to 2016 to present**

Ovid and CINAHL do not have a subject heading for visual hallucinations specifically. Instead, they have one for hallucinations which would result in less specific results so was not used in the search strategy.

Appendix 1.3: Overview of Measures included in the systematic review.

Name of measure	Purpose of measure	Mode of completion	Definition of VH	No. of VH items (overall items included)	Content of VH items	Structure: No. of scales or subsections (no. of items)	Scoring
Questionnaire for Psychotic Experiences (QPE), Schutte et al (2019)	To measure psychotic experiences trans-diagnostically	Scores based on the response of the individual and informants (family/carers)	VHs defined as things seen that others could not see.	15 (50)	Screening question; seen objects, persons, or images that others cannot. Inclusion of questions on frequency, content, onset, relation to unpleasant past experiences, duration, emotional impact, associated distress, functional impact, repetition, location, time, insight, interaction with VHs and frequency of visual illusions.	4 subscales: AH (15), VH (15), Hallucinations in other modalities (6), Delusions (9)	No total score provided; items scored by severity on a 4-to-6-point Likert scale. Subscales = ∑ (items for each subscale)
Tottori University Hallucination Rating Scale (TUHARS), Wada-Isoe et al (2008)	To assess hallucinations in Parkinson's Disease	Clinician-administered interview conducted with the patient and carer.	No definition given.	5 general questions which can be applied to VHs, if these are reported.	Questions evaluate the type of hallucination modality. The VH question asks if the individual has seen something (people, animals, objects) not seen by others. Includes items assessing frequency, severity, caregiver burden, level of insight and psychiatric status at nighttime.	1 scale (5)	Total score= ∑ (all items). Rated on a 4-point Likert scale.

Name of measure	Purpose of measure	Mode of completion	Definition of VH	No. of VH items (overall items included)	Content of VH items	Structure: No. of scales or subsections (no. of items)	Scoring
Psycho-sensory Hallucinations Scale (PSAS), de Chazeron et al (2015)	To measure multi-modal hallucination modalities and their severity	Structured interview	Broad definition of hallucinations as "false or distorted sensory experiences that appear to be real perceptions." No VH specific definition given.	29 (98)	Questions explore the characteristics of reported VHs (e.g., objects, face, animals, people, senses, colour, size), frequency, duration, perception, negative aspects, conviction, impact, and control.	4 subscales: AH (30), VH (29), Olfactory and Gustatory Hallucinations (19), Cenesthetic Hallucinations (19)	No total score provided. Subscales = ∑ (items for each subscale). Binary responses (yes or no) and 4-point Likert scale for frequency.
North East Visual Hallucinations Interview (NEVHI) , Mosimann et al, (2008)	To assess visual hallucinations in older people with cognitive impairment and eye disease.	Clinician rated based semi-structured interview	Defined VHs as perceptions which occur in the absence of a visual stimulus. Reports the distinction between simple and complex VHs.	20 (all VH)	Screening items include open ended questions exploring the things people see that others cannot and the content of such visions (colour, shape, form, movement). Frequency, last occurrence, and onset of all VHs is then assessed. Appraisal, control, insight and acting out of VHs is also explored.	3 subscales: Phenomenology (7), Temporal aspects of hallucinations (4), Emotions, cognitions, and behaviours (9)	No total score provided. Binary responses (yes or no) with follow- up qualitative questions. 5–6-point Likert scale.
The Multi-Modality Unusual Sensory Experiences Questionnaire (MUSEQ), Mitchell et al (2017).	To assess USE in 6 sensory domains: AH, VH, OH, GH, BS, and SP.	Self-report	USE defined as a range of phenomena, e.g., hallucinations and misperceptions where there is discrepancy between what is perceived and what exists. No specific VH definition given.	8 (43)	Assesses the content of VHs, transformation of objects, passing of objects/people, lights/colours appearing more intense.	6 subscales: AH (7), VH (8), OH (8), GH (8), BS (8), SP (4)	Total score= ∑ (all items). Subscales = ∑ (items in each subscale). 5-point Likert scale for frequency.

Name of measure	Purpose of measure	Mode of completion	Definition of VH	No. of VH items (overall items included)	Content of VH items	Structure: No. of scales or subsections (no. of items)	Scoring
The University of Miami Parkinson's Disease Hallucination Questionnaire (UM-PHQ), Papapetropoulos et al, (2008)	To assess hallucinations in Parkinson's Disease in clinical and research contexts.	Clinician rated based on a semi-structured interview	No definition given.	6 which apply to VHs as well as other hallucination types. 9 additional specific to VH. (20)	Quantitative items: modality, frequency, duration, insight, emotional burden. Qualitative items: characteristics; multi-modal hallucinations, time of day, movement, appearance, and the onset of VHs.	2 subscales: Quantitative (6) and Qualitative (14)	Total score= ∑ (all section 1 items). Binary responses of yes/no. Likert scales of varying points.
The Plausibility of Hallucinations Scale (PoH), Franceschi (2011)	To assess the plausibility of hallucinations in Schizophrenia	Not specified.	DSM-IV definition given of hallucinations but not of VH specifically.	7 for VH only and 15 multi-model hallucinations that include VH (50)	Questions include seeing formed VHs, ordinary objects, visions in colour, VH that involve surroundings, bilateral hallucinations, associated with thoughts, feeling, and behaviour, experience of scenes or sequences of animated images. Questions also cover all possible multi-modal hallucinations.	9 sections: Unimodal Hallucinations-AS (10), VH (7), OH (3), TH (2), GH (2), BH (10), Trimodal Hallucinations (10), Quadri-modal Hallucinations (5), Quinti-modal Hallucinations (1)	Binary responses: yes/no, each "yes" response weighting 2 points. Total score =∑ (all items answered "yes" to).
Queen Square Visual Hallucinations Interview (QSVHI), Williams et al (2008)	To assess the occurrence of VH in older people with PD	Clinician rated based on a semi-structured interview	No definition given.	9 (10)	Minor hallucinations e.g., the vivid sensation of a presence in the room, brief vision of movement e.g., a person or animal or looking at something that then appeared as something else (e.g., spots on the wall appearing as insects). Formed VH questions: visions of people, animals, or objects and whether they made any noise. Onset of visual experiences and if they are related to medication or delirium.	5 sections: Screening question to assess the presence of VHs in the past 3 months (1); Minor hallucinations/illusions (3); Formed VH (2), AH (1), Details (3)	No total score provided. Binary responses: yes/no

Name of measure	Purpose of measure	Mode of completion	Definition of VH	No. of VH items (overall items included)	Content of VH items	Structure: No. of scales or subsections (no. of items)	Scoring
Sensory Experiences of Deceased Spouse (SED), Kamp et al (2023)	To assess sensory and quasi-sensory experiences of a deceased spouse.	Self-report	Defined VHs in the context of the deceased as seeing a dead spouse when awake. These experiences do not refer to having vivid dreams or experiences attained through a clairvoyant or medium.	17 (103)	Questions explore the content of VHs including movement, light, colour, transparency, extent to which they appear life-like and have a distinctive appearance and frequency.	11 sections: Presence of sensory experiences of a deceased spouse (5), VH (17), AH (12), TH (7), SP (6), OH and GH (3), Frequency of sensory experiences (4), Relationships with sensory experiences (13), Others' reactions (6), Circumstances during which the sensory experiences occurred (14), Appraisal (16)	No total score provided. Mix of binary responses (yes or no), qualitative questions and 3–6-point Likert scale.
Psychosis and Hallucinations Questionnaire (PsycH-Q), Shine et al (2015)	To assess the presence of hallucinatory symptoms in PD	Self-report	No definition given.	6 (20)	VH questions: presence hallucinations reported as a stimulus moving past in the peripheral field, passage hallucinations which describe a sense of “something” perceived out of the corner of the eye, mistaking an object for something else and the content of VH (animals, objects, people)	Hallucinations & Psychosis (Section 1) includes 3 subscales: VH (6), Misperceptions & hallucinations in other sensory modalities (4), Thought disorder & Psychotic Behaviour (3). Hallucinations Phenotype (Section 2) includes 2 subscales, Attentional Dysfunction (4) & Sleep Impairment (2)	Total score= ∑ (frequency score for Section I, maximum score: 52, and II, maximum score: 28). Subscales = ∑ (items in each subscale). 5-point Likert scales
Psychotic Symptom Rating Scale – Visual Hallucinations (PSYRATS-VH) , Dudley et al (2013)	To assess symptom, change in VHs over the course of therapy for people with VHs.	Clinician rated based on a semi-structured interview	No definition given.	12 (18)	Questions include frequency, duration, location, loudness, cause, amount and degree of negative content and distress, disruption of VH, control and number experienced in the last week.	2 subscales : A: VH (12), B: Delusions (6)	No total score provided. Subscales = ∑ (items in each subscale). 5-point Likert scale

Name of measure	Purpose of measure	Mode of completion	Definition of VH	No. of VH items (overall items included)	Content of VH items	Structure: No. of scales or subsections (no. of items)	Scoring
Multi-Modal Hallucinations Interview (MMHI), Dudley et al (2023)	To assess the presence of multi-sensory/modal hallucinations	Semi-structured interview	No definition given.	9 general questions on USE which can be applied to VH (10)	Questions include whether the VHs and other hallucinations occur simultaneously or serially and if they come from the same source. Distress, content, and belief in the VH are explored.	1 scale (10)	No total score provided. Binary (yes or no) and qualitative responses.
Appraisals and Reactions to Visual Hallucinations Interview (ARVHI), Dudley et al (2012)	To assess appraisals and safety seeking behaviours associated with distressing VH in psychosis	Semi-structured interview	No definition given.	74 (all VH)	Questions explore multi-modal hallucinations, familiarity, and triggers. Affect rates a range of emotions associated with the presence of VH. Appraisals rate VH across a range of beliefs. Responses rates methods under distraction, focus and meta-cognitive approaches. Safety Seeking Behaviours considers the intention and consequences of these.	4 sections: Content of VH (3), Affect associated with VH (3), Appraisals (17), Responses to VH (27), Phenomenology (24)	10-point Likert scale.
Semi- structured interview about visions in psychiatric patients (SSIVPP), Gauntlett-Gilbert & Kuipers (2005)	To assess appraisals in VH.	Clinician rated based on a semi-structured interview.	No definition given.	54 (all VH)	Phenomenology: content, duration, frequency, and beliefs. Triggers: states of mind, environmental impact, and physical sensations. Other related phenomena: multi-modal hallucinations, vivid intrusive imagery, dreams, trauma memories, and other visual experiences. History: VH onset and progression. Content includes the meaning, familiarity, and personality of VH. Cognitions and affect: the impact of VH on cognitions, feelings, behaviours, and personal identity. Control and coping: coping strategies and their success.	7 sections: Phenomenology (9), Triggers (5), Other related phenomena (5), History (3), Content (10), Cognitions and affect (16), Control and coping (6)	5-point Likert scale.

Name of measure	Purpose of measure	Mode of completion	Definition of VH	No. of VH items (overall items included)	Content of VH items	Structure: No. of scales or subsections (no. of items)	Scoring
Rush Hallucination Inventory (RHI) , Pappert et al (1999)	To assess the sleep disturbance, delusions, and hallucinations in people with Parkinson's disease (PD)	Self-report	No definition given.	13 (62)	Questions explore frequency, time of occurrence, duration, situations of occurrence, whether they were frightening for both visual illusions and VH.	4 sections: sleep habits and dreams (11); Visual, auditory, and tactile illusions (18), Hallucinations in different modalities (24), Delusions (9).	No total score provided. 4-point Likert scale or frequency questions with binary responses (yes/no).
Semi-structured interview of Complex Visual Hallucinations (SICVH) , Lai et al (2016)	To assess appraisals of Complex Visual Hallucinations (CVH) and distress in older, non-psychotic people	Clinician rated based on a semi-structured interview.	VH defined as things that some people see that others cannot. These experiences do not refer to dreams, daydreams, or your imagination.	66 (all VH)	Questions explore the presence of VH including the feeling of a presence, seeing multiple copies of something; quantification of VH including onset, length, frequency, time of day, VH number; VH Content (people, children, animals, dots, flashes, motion), familiarity of the VH and whether it speaks; range of emotions associated to VH; perceived control over VH; Insight into what caused the VH, the role it plays in the individual's life, what it means to the person; Appraisals (benevolent and malevolent), power and control of the VH.	7 sections: Presence of VH (7), Quantification of VH (6), Content of VH (16), Affect (14), Perceived control (5), Insight (5), Appraisals (13)	No total score provided. Binary responses (yes, no, don't know) and 3–10-point Likert scales.
Visual Hallucination Questionnaire (VHQ), Van Ommen et al (2019)	assess lifetime '	Clinician administered questionnaire	VH defined as visions when awake or seeing things that other people cannot.	50 (all VH)	Questions explore VH characteristics including onset, time of occurrence, frequency, duration, motion, transparency, colour, size, distress. Questions also assess the presence of simple VH (dots, flashes, letters, words, patterns) and CVH (figures, groups of people, faces with no bodies, animals).	1 scale (50)	No total score provided. Binary responses (yes or no) and Likert scales of varying points.

Note: ∑=sum of; USE= Unusual Sensory Experiences; AH=Auditory Hallucinations; VH=Visual Hallucinations; OF=Olfactory Hallucinations; GH=Gustatory Hallucinations; TH=Tactile Hallucinations; BH= Bimodal Hallucinations; BS=Bodily Sensations; SP=Sensed Presence

Appendix 1.4: Updated COSMIN criteria for good measurement properties

Taken from (Mokkink et al, 2010).

Measurement property	Rating[1]	Criteria
Structural validity	+	**CTT:** CFA: CFI or TLI or comparable measure >0.95 OR RMSEA <0.06 OR SRMR <0.08[2] **IRT/Rasch**: No violation of unidimensionality[3]: CFI or TLI or comparable measure >0.95 OR RMSEA <0.06 OR SRMR <0.08 *AND* no violation of local independence: residual correlations among the items after controlling for the dominant factor < 0.20 OR Q3's < 0.37 *AND* no violation of monotonicity: adequate looking graphs OR item scalability >0.30 *AND* adequate model fit: IRT: χ^2 >0.01 Rasch: infit and outfit mean squares ≥ 0.5 and ≤ 1.5 OR Z-standardized values > -2 and <2
	?	CTT: Not all information for '+' reported IRT/Rasch: Model fit not reported
	–	Criteria for '+' not met
Internal consistency	+	At least low evidence[4] for sufficient structural validity[5] AND Cronbach's alpha(s) ≥ 0.70 for each unidimensional scale or subscale[6]
	?	Criteria for "At least low evidence[4] for sufficient structural validity[5]" not met
	–	At least low evidence[4] for sufficient structural validity[5] AND Cronbach's alpha(s) < 0.70 for each unidimensional scale or subscale[6]

Reliability	+	ICC or weighted Kappa ≥ 0.70
	?	ICC or weighted Kappa not reported
	-	ICC or weighted Kappa < 0.70
Measurement error	+	SDC or LoA < MIC[5]
	?	MIC not defined
	-	SDC or LoA > MIC[5]
Hypotheses testing for construct validity	+	The result is in accordance with the hypothesis[7]
	?	No hypothesis defined (by the review team)
	-	The result is not in accordance with the hypothesis[7]
Cross-cultural validity\measurement invariance	+	No important differences found between group factors (such as age, gender, language) in multiple group factor analysis OR no important DIF for group factors (McFadden's R^2 < 0.02)
	?	No multiple group factor analysis OR DIF analysis performed
	-	Important differences between group factors OR DIF was found
Criterion validity	+	Correlation with gold standard ≥ 0.70 OR AUC ≥ 0.70
	?	Not all information for '+' reported
	-	Correlation with gold standard < 0.70 OR AUC < 0.70
Responsiveness	+	The result is in accordance with the hypothesis[7] OR AUC ≥ 0.70
	?	No hypothesis defined (by the review team)
	-	The result is not in accordance with the hypothesis[7] OR AUC < 0.70

The criteria are based on e.g. Terwee et al.(30) and Prinsen et al.(5)

AUC = area under the curve, CFA = confirmatory factor analysis, CFI = comparative fit index, CTT = classical test theory, DIF = differential item functioning, ICC = intraclass correlation coefficient, IRT = item response theory, LoA = limits of agreement, MIC = minimal important change, RMSEA: Root Mean Square Error of Approximation, SEM = Standard Error of Measurement, SDC = smallest detectable change, SRMR: Standardized Root Mean Residuals, TLI = Tucker-Lewis index

[1] "+" = sufficient, " –" = insufficient, "?" = indeterminate
[2] To rate the quality of the summary score, the factor structures should be equal across studies
[3] unidimensionality refers to a factor analysis per subscale, while structural validity refers to a factor analysis of a (multidimensional) patient-reported outcome measure
[4] As defined by grading the evidence according to the GRADE approach
[5] This evidence may come from different studies
[6] The criteria 'Cronbach alpha < 0.95' was deleted, as this is relevant in the development phase of a PROM and not when evaluating an existing PROM.
[7] The results of all studies should be taken together and it should then be decided if 75% of the results are in accordance with the hypotheses

Appendix 1.5: Measures excluded from Aynsworth et al., (2017) review.

Name of Measure, author(s), year	Reason for exclusion
General measure of psychotic (like) symptoms	
The Oxford-Liverpool Inventory of Feelings and Experiences (O-LIFE), Mason et al (1995)	Only 3 items on visual hallucinations (VH); no separate VH scale/subscale.
Rating Scale for Psychotic Experience (RSPS), Chouinard & Miller (1999)	3 VH items; no separate VH scale/subscale.
The Cardiff Anomalous Perceptions Scale (CAPS), Bell et al (2006)	No separate VH scale/subscale.
Parkinson disease-associated psychotic symptoms questionnaire (RSPS), Fenelon et al. (2010)	4 VH items; no separate VH scale/subscale.
General Hallucination Measures	
Launay Slade Hallucination Scale - Revised (LSHS-Ra), Morrison et al (2000)	4 VH items; no separate VH scale/subscale.
Launay Slade Hallucination Scale – Revised (LSHS-Rb), Laroi et al (2004)	2 VH items; no separate VH scale/subscale.
Revised Hallucination Scale (RHS), Morrison et al (2002)	5 VH items; no separate VH scale/subscale.
Specific Visual Hallucination Measures	
Institute of Psychiatry Visual Hallucination Interview (IPVHI), Santhouse et al (2000)	Developed to assess the phenomenology of VH in eye disease.

Appendix 1.6: COSMIN Criteria for sufficient content validity

Taken from (Terwee et al., 2018), pp.8.

Ten criteria for good content validity

Relevance	
1	Are the included items relevant for the construct of interest?
2	Are the included items relevant for the target population of interest?
3	Are the included items relevant for the context of use of interest?
4	Are the response options appropriate?
5	Is the recall period appropriate?
Comprehensiveness	
6	Are no key concepts missing?
Comprehensibility	
7	Are the PROM instructions understood by the population of interest as intended?
8	Are the PROM items and response options understood by the population of interest as intended?
9	Are the PROM items appropriately worded?
10	Do the response options match the question?

Appendix 1.7: Overall Content Validity and Quality of Evidence Ratings

Measure	COSMIN Methodological Quality Rating of Measure's Development Study	Overall Quality of the Evidence Rating	Rationale
ARVHI	Inadequate	*Very Low*	-3 only development study of inadequate quality
MMHI	Inadequate	Very Low	-3 only development study of inadequate quality
MUSEQ	Inadequate	Very Low	-3 only development study of inadequate quality
NEVHI	Doubtful	Low	-2 only development study of doubtful quality
PoH	Inadequate	Very Low	-3 only development study of inadequate quality
PSAS	Inadequate	Very Low	-3 only development study of inadequate quality
PsycH-Q	Inadequate	Very Low	-3 only development study of inadequate quality
PSYRATS-VH	Inadequate	Very Low	-3 only development study of inadequate quality
QPE	Inadequate	Very Low	-3 only development study of inadequate quality
QSVHI	Inadequate	Very Low	-3 only development study of inadequate quality
RHI	Inadequate	Very Low	-3 only development study of inadequate quality
SSIVPP	Inadequate	Very Low	-3 only development study of inadequate quality
SICVH	Inadequate	Very Low	-3 only development study of inadequate quality
SED	Doubtful	Low	-2 only development study of doubtful quality
TUHARS	Inadequate	Very Low	-3 only development study of inadequate quality
UM-PHQ	Inadequate	Very Low	-3 only development study of inadequate quality
VHQ	Inadequate	Very Low	-3 only development study of inadequate quality

Appendix 1.8: Psychometric properties reported in the included articles.

Measure	Country (language) in which the measure was evaluated	Structural validity		Internal consistency		Cross-cultural validity\ measurement invariance	Reliability	
		N	Result	N	Result (rating)		N	Result (rating)
QPE, Rossell et al (2019)	Australia (English)	173	Current VH subscale (0.97), Lifetime VH subscale (CFI: 0.99) Current AH subscale (CFI 0.99), Lifetime AH subscale (0.93), 1 Unidimensional model present for delusion subscale (Current delusions, CFI, 0.97, lifetime 0.9) (+)	173	Cronb. Alpha =0.75-0.90 for lifetime experiences, 0.70-0.89 for current experience (+)	Not Reported	16 (NB. 14 for VH ICC analysis)	Test-retest ICC = 0.70-0.92. (+) Inter-rater ICC =0.99-1.00 (+)
QPE-Arabic Translation, Yehra et al (2023)	Quatar (Arabic)	Not reported		50	Cronb. Alpha = 0.75 (AHs), 0.74 (VHs), 0.79 (Delusions) (+)	Not reported	5 37	Inter-rater ICC=0.83-1.00 (+) Test-retest ICC=0.71-0.91 (+) Current THs ICC=0.53 (-) Current VHs ICC=0.68 (-)
Pooled or summary result (overall rating)		**173**	**0.90-0.99 (+)**	**223**	**0.70-0.90 (+)**		**58**	**0.53-1.00 (-)**

NEVHI, Mosimann et al (2008)	UK (English)	114	PCA varimax rotation: 4 factor model 0.835-0.969 (+)	114	Cronb. Alpha =0.71 for VH scale (+)	Not reported	114	Inter-rater K=0.72 (simple VH), K=0.83 (complex VH) (+)
NEVHI validity study in PD population, Holiday et al (2017)	America (English)	Not reported						
Pooled or summary result (overall rating)		**114**	**0.835-0.969 (+)**	**114**	**0.71 (+)**		**114**	**0.72-0.83 (+)**
MUSEQ, Mitchell et al (2017)	Australia (English)	513	Correlated Factors Model: CFI=0.95 Bi-factor Model: CFI=0.95 (+)	513	Cronb. Alpha =0.82 (AH), 0.88. (VH), 0.87 (OH), 0.88 (GH), 0.88 (BS), 0.77 (SP) (+)	Not reported	96	Test-retest ICC =0.56-0.70 (-)
PSAS, de Chazeron et al (2015)	France (French)	137 (86=PD patients,51 =SCZ patients)	EFA: 4 factors accounted for 61% of variance for PD patients, 3 factors accounted for 55% variance for SCZ patients (?)	137	Kuder-Richardson Alpha coefficient: PD Group: 0.49-0.77 (VH subscale =0.49) (-) SCZ Group: 0.56-0.71 (VH subscale =0.65 (-)	Not reported	Test-retest: 57 (37 PD, 20 SCZ) Inter-rater: 35 (17 PD, 18 SCZ)	Inter-rater K=0.62-0.88(PD Group), 0.56-0.86 (SCZ Group). On both PD and SCZ, VH subscale had the lowest scores. (-) Test-retest K =0.62-0.87 (PD Group), K=0.48-0.80 (SCZ Group) (-)
PsycH-Q, Shine et al (2014)	Australia (English)	Not reported		197	Cronb. Alpha =0.70-0.92 (?)	Not reported	44	Test-retest ICC=0.93 (+)
TUHARS, Wada-Isoe et al (2008)	Japan (Japanese)	Not reported		41	Cronb. Alpha=0.88 (?)	Not reported	Not reported	

Measure	Country (language) in which the measure was evaluated	Measurement error	Criterion validity	Hypotheses testing	Responsiveness
		N	Result (rating)	N	Result (rating)
QPE, Rossell et al (2019)	Australia (English)	Not reported	173	SAPS VH lifetime subscale significantly correlated with QPE VH lifetime subscale (0.68) (-)	173
QPE-Arabic Translation, Yehra et al (2023)	Quatar (Arabic)	Not reported	46	PANSS Hallucination subscale significantly correlate with QPE VH lifetime (0.45) and current (0.58) (-)	46
Pooled or summary result (overall rating)		**173**	**0.58-0.68 (-)**	**219**	**2+ and 1- overall (?)**
NEVHI, Mosimann et al (2008)	UK (English)	Not reported	Not reported	Not reported	Not reported
NEVHI validity study in PD population, Holiday et al (2017)	America (English)	Not reported	Not reported	117	Convergent validity: Results in line with 3 hypotheses (3 +)
Pooled or summary result (overall rating)			117 (*114 in one correlation with visual acuity measure)	**3+ overall (+)**	
MUSEQ, Mitchell et al (2017)	Australia (English)	Not reported	Not reported	513	Convergent validity: Results in line with 3 hypotheses (3+) Discriminant validity: Results in line with 1 hypothesis (1+) (4+)

Appendix 1.9 Overall Psychometric Quality Ratings and Rationale

Measure	Psychometric Property	COSMIN Methodological Quality Rating	Overall Quality of the Evidence Rating	Rationale
QPE	SV	Adequate	Moderate	-1 study of adequate quality
	IC	Very Good	High	N/A
	R	Doubtful	Low	-1 Two studies of doubtful quality -1 Sample size 50-100
	CV	Very Good	High	N/A
	HT	Adequate	High	N/A
NEVHI	SV	Adequate	Moderate	-1 study of adequate quality
	IC	Very Good	High	N/A
	R	Doubtful	Low	-2 study of doubtful quality
	HT	Doubtful	Low	-2 study of doubtful quality
MUSEQ	SV	Very Good	High	N/A
	IC	Very Good	High	N/A
	R	Adequate	Low	-1 study of adequate quality -1 sample size 50-100
	HT	Very Good	High	N/A

Measure	Psychometric Property	COSMIN Methodological Quality Rating	Overall Quality of the Evidence Rating	Rationale
PSAS	SV	Inadequate	Very low	-3 study of inadequate quality
	IC	Very Good	High	N/A
	R	Adequate	Low	-1 study of adequate quality -1 sample size 50-100
	CV	Very good	High	N/A
PsycH-Q	IC	Very Good	High	N/A
	R	Doubtful	Very low	-2 study of doubtful quality -2 sample size <50
	HT	Adequate	Moderate	-1 study of adequate quality
	CV	Very Good	Moderate	-1 sample size 50-100
TUHARS	IC	Very Good	Low	-2 sample size <50
	HT	Adequate	Very Low	-1 study of adequate quality -2 sample size <50

Note: SV=Structural Validity; IC=Internal Consistency; R=Reliability; CV=Criterion Validity; HT=Hypothesis Testing (for construct validity)

Appendix 2.1: MRP Proposal

The proposal for this MRP is available at:

OSF | Exploring the phenomenology of hallucinatory experiences and their relation to trauma re-experiencing symptoms: The Questionnaire of Hallucinatory Experiences

Document title: Proposal V5

Appendix 2.2 Revised Analysis Plan

The revised analysis plan for this MRP is available at:

OSF | Exploring the phenomenology of hallucinatory experiences and their relation to trauma re-experiencing symptoms: The Questionnaire of Hallucinatory Experiences

Document title: Revised Analysis plan, June 2023

Appendix 2.3 Ethical Approval

University of Glasgow

Professor Andrew Gumley

MVLS College Ethics Committee

Developing a Self-report Evaluation of Hallucinatory Experiences: The SEHE survey

200210004

The College Ethics Committee has reviewed your application and has agreed that there is no objection on ethical grounds to the proposed study. We are happy therefore to approve the project, subject to the following conditions

- Project end date as stipulated in original application.
- The data should be held securely for a period of ten years after the completion of the research project, or for longer if specified by the research funder or sponsor, in accordance with the University's Code of Good Practice in Research: (http://www.gla.ac.uk/media/media_227599_en.pdf)
- The research should be carried out only on the sites, and/or groups defined in the application.
- Where a questionnaire collects sensitive or protected information, the Data Protection Impact Assessment should be discussed with the data office.
- Any proposed changes in the protocol should be submitted for reassessment, except when it is necessary to change the protocol to eliminate hazard to the subjects or where the change involves only the administrative aspects of the project. The Ethics Committee should be informed of any such changes.
- For projects requiring the use of an online questionnaire, the University has an Online Surveys account for research. To request access, see the University's application procedure at https://www.gla.ac.uk/research/strategy/ourpolicies/useofonlinesurveystoolforresearch/.
- You should submit a short end of study report within 3 months of completion.

Yours sincerely

Terry Quinn
FWSO, FESO, MD, FRCP, BSc (hons), MBChB (hons)
Reader / Honorary Consultant

Appendix 2.4 Questionnaire of Hallucinatory Experiences (QHE)

The QHE is available at:

OSF | Exploring the phenomenology of hallucinatory experiences and their relation to trauma re-experiencing symptoms: The Questionnaire of Hallucinatory Experiences

Document title: The Questionnaire of Hallucinatory Experiences

Appendix 2.5 Consent Form

The consent form for the study is available at:

OSF | Exploring the phenomenology of hallucinatory experiences and their relation to trauma re-experiencing symptoms: The Questionnaire of Hallucinatory Experiences

Document title: Consent Form, Phase 3, V2

Appendix 2.6 Participant Information Sheet

The Participant Information Sheet (PIS) is available at:

OSF | Exploring the phenomenology of hallucinatory experiences and their relation to trauma re-experiencing symptoms: The Questionnaire of Hallucinatory Experiences

Document title: Participant Information Sheet