

University of Groningen

## The development, validity, and reliability of the auditory vocal hallucination rating scale questionnaire (AVHRS-Q)

Steenhuis, Laura A; Pijnenborg, Gerdina H M; Visser, Ellen; van de Willige, Gerard; van Beilen, Marije; Nauta, Maaïke H; Aleman, André; Bartels-Velthuis, Agna A

*Published in:*  
Social Psychiatry and Psychiatric Epidemiology

*DOI:*  
[10.1007/s00127-019-01692-z](https://doi.org/10.1007/s00127-019-01692-z)

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2019

[Link to publication in University of Groningen/UMCG research database](#)

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

Steenhuis, L. A., Pijnenborg, G. H. M., Visser, E., van de Willige, G., van Beilen, M., Nauta, M. H., Aleman, A., & Bartels-Velthuis, A. A. (2019). The development, validity, and reliability of the auditory vocal hallucination rating scale questionnaire (AVHRS-Q). *Social Psychiatry and Psychiatric Epidemiology*, 54(8), 927-935. <https://doi.org/10.1007/s00127-019-01692-z>

### **Copyright**

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

### **Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*



# The development, validity, and reliability of the auditory vocal hallucination rating scale questionnaire (AVHRS-Q)

Laura A. Steenhuis<sup>1</sup> · Gerdina H. M. Pijnenborg<sup>1,2</sup> · Ellen Visser<sup>3</sup> · Gerard van de Willige<sup>3</sup> · Marije van Beilen<sup>4</sup> · Maaike H. Nauta<sup>1</sup> · André Aleman<sup>1,5</sup> · Agna A. Bartels-Velthuis<sup>3</sup>

Received: 17 October 2018 / Accepted: 9 March 2019 / Published online: 22 March 2019  
© The Author(s) 2019

## Abstract

**Purpose** The Auditory Vocal Hallucination Rating Scale Questionnaire (AVHRS-Q) is a short self-report measure assessing several characteristics of auditory vocal hallucinations (AVH) that was derived from a validated clinical interview (the auditory vocal hallucination rating scale; AVHRS). This study investigated the internal reliability, convergent validity, and divergent validity of the AVHRS-Q using two clinical samples.

**Methods** In sample I, 32 psychiatric patients with AVH were recruited from an academic hospital service and assessed with the AVHRS and the AVHRS-Q. Data for sample II were retrospectively retrieved from a pseudonymised Routine Outcome Monitoring (ROM) database collected in the context of mental healthcare at the same academic hospital service. Data from 82 psychiatric patients with AVH were retrieved, who completed the AVHRS-Q, and measures of psychological distress (the Outcome Questionnaire; OQ-45, and the Symptom Checklist; SCL-90) and quality of life (the Manchester Short Assessment of Quality of Life; MANSA).

**Results** The AVHRS-Q showed good internal consistency in both samples. Severity scores of the AVHRS-Q were strongly correlated to the severity scores of the AVHRS ( $r=0.90$ ,  $p<0.01$ ). The AVHRS-Q and AVHRS did not differ in the identification of mild and severe voice-hearers [ $\chi^2(1, N=32)=15.71$ ]. AVHRS-Q severity scores had moderate correlations with measures of psychological distress (OQ-45,  $r=0.43$ ,  $p<0.01$ ; SCL-90,  $r=0.50$ ,  $p<0.05$ ) and quality of life (MANSA,  $r=-0.22$ ,  $p<0.01$ ).

**Conclusions** The AVHRS-Q demonstrated good reliability, convergent validity, and divergent validity, suggesting it can be applied in both clinical and research settings for a quick and reliable assessment of AVH.

**Keywords** Questionnaire · Validation study · Hallucinations · Psychosis

## Introduction

Auditory vocal hallucinations (AVH) are prevalent in children, adolescents, and adults, both in clinical settings and in the general population [1–4]. AVH are common in psychotic

illnesses and other mental disorders such as depression, bipolar, dissociative, and substance use disorders [5]. AVH severity may be predictive of (amongst others) social problems [6, 7], suicidal ideation [8] or substance abuse [9] and a reliable assessment of AVH is therefore very important.

✉ Laura A. Steenhuis  
l.a.steenhuis@rug.nl

<sup>1</sup> Department of Clinical Psychology and Experimental Psychopathology, Faculty of Behavioural and Social Sciences, University of Groningen, Grote Kruisstraat 2/1, 9712 TS Groningen, The Netherlands

<sup>2</sup> Department of Psychotic Disorders, GGZ-Drenthe, Dennenweg 9, 9404 LA Assen, The Netherlands

<sup>3</sup> University Medical Center Groningen, University Center for Psychiatry, Rob Giel Research center, University

of Groningen, Hanzeplein 1 (CC72), 9713 GZ Groningen, The Netherlands

<sup>4</sup> University Medical Center Groningen, University Center for Psychiatry, Psychosis Department, University of Groningen, Hanzeplein 1, 9713 GZ Groningen, The Netherlands

<sup>5</sup> Department of Neuroscience, University of Groningen, University Medical Center Groningen, Antonius Deusinglaan 1, 9713 AV Groningen, The Netherlands

Given that voice-hearing is an internal experience which cannot be directly observed or measured, investigating AVH relies on the report of individual experiences. The most reliable manner to do this is using structured interviews and self-report instruments.

In 2012, Bartels-Velthuis and colleagues validated the AVHRS, a structured interview to gain insight into the characteristics of voices [10] and from which a severity measure of voice-hearing can be derived. The AVHRS distinguishes itself from other measures for AVH, as besides the qualitative characteristics and severity of AVH, it also assesses the form and content of voices (in contrast to the BAVQ-R; [11]) and the number of voices (in contrast to the PSYRATS; [12]) (see the validation paper [10] for a more elaborate description). Given that there has been a shift from interview measures to self-report measures of AVH [13], a questionnaire version of the AVHRS was warranted and has now been developed. Indeed, self-report measures have the benefit of being inexpensive and time-efficient and do not require training of assessors. This is especially useful when quick or frequent assessment of AVH is required. For example, in clinical practice this may be necessary for routine outcome monitoring (ROM) or for clinical intakes. Self-report measures may also be useful for research on clinical therapies, to examine to what extent or at which time-point, the occurrence, characteristics, and severity of AVH are changing.

A number of self-report measures for AVH are available [13]. These questionnaires are usually tailored to measure a specific aspect of AVH, for example beliefs about AVH (e.g., BAVQ-R; [11]), interpretations and attitudes towards AVH (e.g., VPD; [14]), coping with AVH (RAHQ; [15]), and mindfulness of AVH (SMVQ; [16]). There are some questionnaires on AVH that have a wider focus and are also quite brief (13 items, Hamilton Program for Schizophrenia Voices Questionnaire (HPSVQ), [17]; ten items, the delusion and voices self-assessment (DV-SA), [18]). However, these questionnaires do not incorporate items on the form of address (1st, 2nd or 3rd person), the location of voices,

separate or simultaneous voices, severity of negative content, or whether the voices make them anxious. The DV-SA specifically does not enquire about the duration or loudness of voices, or whether negative voices are present. Overall, compared to previous measures the AVHRS-Q ensures a comprehensive assessment of AVH, encompassing multiple qualitative aspects of AVH (e.g., negative voices, distress, interference with thinking and daily functioning) in a set of 17 items.

The aim of this study is to validate a self-report version of the auditory vocal hallucination rating scale (AVHRS; [10, 19]), called the AVHRS-Q(uestionnaire). In this validation study, the internal consistency, convergent validity and divergent validity of the AVHRS-Q will be examined. It is expected that the AVHRS-Q will correlate highly with the interview version (AVHRS; [10]), demonstrating good convergent validity. As greater severity of AVH is related to both increased psychological distress and a lower quality of life [20], it is expected that the severity measure of the AVHRS-Q will correlate with measures of psychological distress (the outcome questionnaire, OQ-45; [21] and the symptom checklist, SCL-90; [22]) and quality of life (the Manchester Short Assessment of Quality of Life, MANSA; [23]). However, given that the AVHRS-Q specifically measures AVH characteristics and severity, and not general psychological distress or quality of life, the correlations between these measures are expected to be no more than moderate, indicative of divergent validity.

## Methods

### Participants and procedures

For the current study data of two clinical samples (for demographics see Table 1) were used. Inclusion criteria consisted of receiving treatment for AVH, being between 18 and 65 years old, and having a good command of the

**Table 1** Demographics of sample I and II

	Sample I ( <i>n</i> = 32)	Sample II ( <i>n</i> = 82)
Female ( <i>n</i> , %)	18 (56)	44 (54)
Age ( <i>M</i> , <i>SD</i> )	38 (11)	39 (12)
Duration of AVH in years ( <i>M</i> , <i>SD</i> )	12 (11)	N.A.
Psychotic disorder <sup>§</sup> ( <i>n</i> , %)	29 (91)	60 (73)
Also diagnosed with a comorbid disorder	8 (28)	24 (40)
Other disorder (e.g., mood disorder, personality disorder, anxiety disorder)	3 (9)	22 (27)
Also diagnosed with a comorbid disorder	0 (0)	8 (36)

AVH auditory vocal hallucinations, N.A. not available

<sup>§</sup>Psychotic disorders in these samples consisted of Schizophrenia, Schizoaffective disorder, and Psychotic disorder not otherwise specified

Dutch language. Exclusion criteria consisted of having an organic brain disorder. Approval for the study with sample I was obtained from the Medical Ethics Committee of the University of Medical Center Groningen (ref: M13.146159). The sample size for sample I was calculated a priori by a statistician. It was determined that at least 31 people were required to obtain a two-sided confidence interval with minimum length of 0.1 for a correlation of 0.9. Thirty-two patients with AVH were recruited for sample I at the Voices Outpatient Department of the University Medical Center Groningen (The Netherlands). Patients were approached by their therapist or by the research coordinator of the Voices Outpatient Department, and received both verbal and written information about the research study, including an informed consent form. Upon providing written informed consent, participants were contacted by the researcher and completed the AVHRS-Q and were interviewed with the AVHRS. During the study, participants alternately started with the self-report version of the AVHRS (AVHRS-Q) or with the interview (AVHRS) to rule out selective memory biases for one of the measurements. Data collection for sample I took place from February 2011 until December 2015.

Data for sample II were retrospectively retrieved from a pseudonymised routine outcome monitoring (ROM) database collected in the context of mental healthcare at the Voices Outpatient Department of the University Medical Center Groningen (The Netherlands). All patients who are referred for treatment to the University Medical Center Groningen take part in ROM assessments and are informed that their data may be used for research purposes whilst having the option to opt-out. Given that these data were collected in the context of treatment and not for research purposes—therefore not requiring the patient to change their behavior for the research—no additional ethical approval for the data is required according to Dutch legislation. Sample II consisted of 82 patients with AVH receiving treatment at the Voices Outpatient Department of the University Medical Center Groningen (The Netherlands). At the start of their treatment, they completed the AVHRS-Q, the MANSA and either the OQ-45 ( $n=62$ ) and/or the SCL-90 ( $n=24$ ) (depending on which instrument their therapist selected) through the ROM service. The requested ROM data were collected from October 2011 until February 2017. As the current study took part in The Netherlands, all questionnaires and interviews were completed in Dutch language.

## Measures

### The AVHRS and development of the AVHRS-Q

The AVHRS [10, 19] is a structured 16-item interview, administered by an experienced therapist to evaluate AVH during a period of 1 month. The AVH are rated on four- and

five-point scales in terms of frequency, duration, loudness, negative content, distress, anxiety, control, and interference with thinking and daily life. Scores range from 0 (not applicable) to 3 or 4 (most applicable).

The AVHRS-Q [24] is the self-report version of the AVHRS, designed to be administered without the presence of an interviewer, therapist or researcher. A full version of the AVHRS-Q can be downloaded at <https://www.rgoc.nl/downloads> (see Table 2 for a summary of the items). The AVHRS-Q has 17 items, 15 of which are assessed with a four- and five-point scale and two on a ten-point scale. For the four- and five-point scales, scores range from 0 (not applicable) to 3 or 4 (most applicable). For the ten-point scales, scores ranging from 1 (not at all/never) to 10 (extremely/always). The items of the AVHRS-Q were based on the items of the AVHRS, but adapted somewhat for the purpose of self-report administration. The first version of the AVHRS-Q was evaluated by ten patients with AVH. Based on their feedback and input from experts in the field, ten questions from the original AVHRS were refined and one item was expanded into two items. To specify, some items of the AVHRS-Q received more answer options in comparison to the interview version (see Table 2, items 3, 4, 7, 8, 10, and 15). For example, the item assessing duration of voices has four answer options in the interview version (seconds, minutes, 1 h, and several hours to continuously) in comparison to five options in the questionnaire (see Table 2, item 4). Moreover, the wording of some items was reformulated to be more simple and unambiguous (see Table 2, item 12 and 13). Additionally, the AVH frequency and intensity of suffering in the AVHRS-Q (see Table 2, item 16 and 17) are rated on a ten-point scale instead of a five-point scale in the AVHRS, as to be more sensitive to subtle changes over time.

In accordance with the AVHRS and previous publications with this measure [4, 7, 25, 26], a severity index can be composed from the individual items of the AVHRS-Q. Items regarding the number of voices, localization of voices and hypnagogic/hypnopompic hallucinations are not included in the severity index (see previous publications; [10]). The answers to individual items are recoded to ‘0’ (none to mild consequences) or ‘1’ (considerable to severe consequences). Subsequently, a sum score of the recoded items is created, ranging from 0 to 14. In addition to the AVHRS-Q providing an overall severity measure of AVH, the individual items can also be used to yield specific information on characteristics of AVH (see Table 2).

### Quality of life

Quality of life was assessed with the Manchester Short Assessment of Quality of Life (MANSA; [23]), a 16-item self-report measure consisting of four objective items and 12 subjective items [satisfaction with life, accommodation,

**Table 2** Summary of individual items of the AVHRS-Q and construction of the severity index

AVHRS-Q item	Options	Severity index (point awarded per option)
1. Number of voices	Always one voice	–
	More than one voice	
2. Separately or simultaneously	Speaking separately	0
	Speaking separately and together	1
	Always speaking together	1
3. Frequency	Not more than once a month	0
	Once a week	0
	Once a day	0
	Once an hour	1
	More than once an hour	1
4. Duration	A few seconds	0
	A few minutes	0
	A few minutes to about 15 min	0
	15 min to about an hour	1
	One hour to almost constantly	1
5. Hypnagogic and/or hypnopompic voices	Voices when falling asleep	–
	Voices when waking up	
	Voices both when falling asleep and when waking up	
	Voices occur at all times	
6. Location	Inside my head	–
	Inside my head and from the outside environment	
	From the outside environment (in the immediate vicinity)	
	From the outside environment, further away	
7. Form of address	The voices say what I am thinking (first person)	0
	The voices speak to me (second person)	0
	The voices speak to me and about me (second and third person)	1
	The voices speak about me (third person)	1
8. Loudness	Soft whisper or murmuring	0
	More quiet than own voice	0
	As loud as own voice	0
	Louder than own voice	1
9. Positive or negative content	Much louder than own voice (yelling, shouting, and screaming)	1
	Always positive	0
	Mostly positive but occasionally negative	0
	Mostly neutral or equally positive and negative	0
	Mostly negative	1
10. Severity of negative content	Always negative	1
	Negative, but not about myself or others	0
	Negative about what I or others do	0
	Negative about how I am or how others are	1
	Threatening commands or orders	1
11. Anxiety	Never	0
	Occasionally	0
	Half of the time	1
	Most of the time	1
	All of the time	1
12. Interference with daily functioning	Never	0
	Occasionally	0
	Half of the time	0

**Table 2** (continued)

AVHRS-Q item	Options	Severity index (point awarded per option)
13. Interference with thoughts	Most of the time	1
	All of the time	1
	Never	0
	Occasionally	0
	Half of the time	1
14. Control	Most of the time	1
	All of the time	1
	Always	0
	Most of the time	0
	Half of the time	0
15. Attribution of origin	Less than half of the time	1
	Never	1
	The voices only concern myself	0
	The voices mostly concern myself	0
	Equally concern myself as outside influences	1
16. Frequency of distress (1 = never, 10 = always)	Mostly caused by outside influences	1
	Fully caused by outside influences	1
	0–6	0
	7–10	1
	17. Intensity of distress (1 = not at all, 10 = extremely)	0–6
7–10		1
Total		Minimum = 0 Maximum = 14

This table represents a summary of the questionnaire and should not be used for administration. The full questionnaire can be downloaded at: <https://www.rgoc.nl/downloads>

AVHRS-Q auditory vocal hallucination rating scale questionnaire

housemates (or living alone), leisure activities, physical health, psychological health, personal safety, friendships, relationship to family, (absence of) romantic relationship, sex life, and financial circumstances]. Items are rated on a seven-point Likert scale, ranging from 1 ‘could not be worse’ to 7 ‘could not be better’. The summary score consists of the mean of the twelve subjective items, with higher scores indicating better quality of life.

### Psychological distress

Psychological distress was assessed with either the outcome questionnaire (OQ-45) or the symptom checklist (SCL-90). Given that the data were collected through ROM assessments for treatment purposes, the therapist was free to choose which assessment measure was administered to the patient on the basis of the therapists’ own preference and familiarity with the instrument. For the current study, both questionnaires were selected, as using only one would have led to a loss of information. The OQ-45 [21] is a 45-item self-report measure assessing clinical outcome in terms of

symptom distress, interpersonal relations and social role performance. For this study, the symptom distress subscale was used consisting of 25 items. Each item is scored on a five-point rating scale, from never ‘0’ to almost always ‘4’. A sum score denoting psychological distress was computed by adding up all items, with high scores pointing to more distress. The SCL-90 [22] is a 90-item self-report measure, assessing a variety of psychopathology. Each item is rated on a five-point rating scale, from ‘1’ (never) to ‘5’ (almost always). The items are clustered in nine dimensions: somatization, obsessive–compulsive, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation, and psychoticism. A sum score denoting psychological distress was computed by adding up all items. Higher scores suggest a lower level of psychological and physical functioning.

### Statistical analyses

Analyses were carried out using SPSS version 23 for Windows [27]. In sample I, two severity groups were created

separately for both the AVHRS and AVHRS-Q: those with ‘severe AVH’ (scoring in the highest quartile of the severity index, i.e., in our study 10 or higher) and with ‘mild AVH’ (scoring 0–9).

To examine convergent validity in sample I, Pearson correlation coefficients between total severity scores and separate items of the AVHRS and AVHRS-Q were computed. A paired-samples *t* test was performed to examine the differences in the mean AVH severity score between the AVHRS-Q and the AVHRS. An exact McNemar’s test was used to examine the distribution of mild and severe AVH groups between the two measures. Internal consistency of both instruments was determined by calculating Cronbach’s alpha [28].

To examine divergent validity in sample II, Pearson correlation coefficients between the total severity score of the AVHRS-Q and total score of the MANSA, the SCL-90 and OQ-45, were computed.

## Results

### Descriptives

In sample I, the AVHRS-Q took an average of 5.8 min to be completed (SD: 2.72, range: 2–15), whereas the AVHRS took an average of 14.3 min to administer (SD: 4.69, range: 8.3–27).

### Internal consistency

In sample I, Cronbach’s alpha of both the AVHRS-Q and the AVHRS was 0.87. In sample II, Cronbach’s alpha of the AVHRS-Q amounted to 0.78.

### Convergent validity

The average severity scores and severity groups for both the AVHRS-Q and AVHRS for sample I are given in Table 3. The severity measures of the AVHRS and AVHRS-Q were highly correlated. This correlation did not differ for

participants who started with the AVHRS ( $r=0.90, p<0.01$ ) and those who started with the AVHRS-Q ( $r=0.89, p<0.01$ ). The Pearson correlation coefficients between individual corresponding items of both measures ranged from 0.44 (moderate) to 0.82 (high), with a median of 0.72 (see Table 4). The mean AVH severity measure and the distribution of severity groups did not differ significantly between the AVHRS-Q and the AVHRS.

**Table 4** Correlations between individual items of the AVHRS-Q and AVHRS

AVHRS-Q item	Pearson correlation with corresponding item on AVHRS
1. Number of voices	0.86**
2. Separately or simultaneously	0.84**
3. Frequency	0.73**
4. Duration	0.70**
5. Hypnagogic and/or hypnopompic voices <sup>§</sup>	
6. Location	0.71**
7. Form of address	0.56**
8. Loudness	0.78**
9. Positive or negative content	0.85**
10. Severity of negative content	0.57**
11. Anxiety	0.74**
12. Interference with daily functioning	0.46**
13. Interference with thoughts	0.52**
14. Control	0.82**
15. Attribution of origin	0.84**
16. Frequency of distress	0.44*
17. Intensity of distress	0.70**

AVHRS-Q auditory vocal hallucination rating scale questionnaire, AVHRS auditory vocal hallucination scale

\* $p<0.05$ , \*\* $p<0.01$

<sup>§</sup>Could not be computed as all participants consistently reported option 4 (at all times of the day)

**Table 3** Average AVH severity score and distribution of severity groups for the AVHRS-Q and AVHRS (sample I,  $N=32$ )

	AVHRS-Q	AVHRS	Pearson’s <i>r</i>	Paired-samples <i>t</i> test	Exact Mc Nemar’s test
AVH severity score ( <i>M</i> , SD)	6.91 (3.15)	7.38 (3.66)	$r=0.90^{**}$	$t(31)=-0.55$	
AVH severity group ( <i>N</i> , %)					
Mild	24 (75%)	22 (68.8%)			$\chi^2(1, N=32)=15.71$
Severe	8 (25%)	10 (31.3%)			

AVH auditory vocal hallucinations, AVHRS-Q auditory vocal hallucination rating scale questionnaire, AVHRS auditory vocal hallucination rating scale

\* $p<0.05$ , \*\* $p<0.01$

**Table 5** Comparison of average AVH severity scores (AVHRS-Q) with measures of quality of life (MANSA) and psychological distress (OQ-45 and SCL-90) (sample II)

	<i>M</i> (SD)	MANSA ( <i>n</i> = 82)	SCL-90 ( <i>n</i> = 24)	OQ-45 ( <i>n</i> = 62)
AVH severity score (Pearson's <i>r</i> )	7.66 (2.69)	−0.22**	0.50*	0.43**

AVH auditory vocal hallucinations, AVHRS-Q auditory vocal hallucination rating scale questionnaire, AVHRS auditory vocal hallucination rating scale, MANSA the manchester short assessment of quality of life, SCL-90 symptom checklist, OQ-45 outcome questionnaire

\* $p < 0.05$ , \*\* $p < 0.01$

### Divergent validity

Descriptives of sample II are given in Table 5. In sample II, the AVHRS-Q severity score was moderately correlated with the psychological distress (OQ-45 and SCL-90) and the quality of life (MANSA) scores. AVH severity was not significantly different between those who did and did not complete the OQ-45 [ $t(80) = 0.46$ ,  $p > 0.05$ ] and SCL-90 [ $t(80) = -0.48$ ,  $p > 0.05$ ].

### Discussion

The current study shows that the auditory vocal hallucination rating scale questionnaire (AVHRS-Q) [24] is a reliable and valid self-report instrument to assess the characteristics and severity of auditory vocal hallucinations (AVH). The findings demonstrate that the AVHRS-Q converges highly with the interview measure on which it was based (the AVHRS; [10]). In addition, the AVHRS-Q is shown to be a specific measure of AVH and not a general measure of psychological distress (OQ-45; [21] and SCL-90; [22]) or quality of life (MANSA; [23]). Internal reliability of the AVHRS-Q was found to be good and comparable to the reliability of the AVHRS.

The AVHRS-Q severity scores correlate highly with the corresponding severity scores of the interview version (AVHRS). In addition, the AVHRS-Q and the AVHRS did not identify a different proportion of patients as having 'mild' or 'severe' AVH. This implies that the already validated and widely used AVHRS [10, 29–31] can now also be used in the self-report version for the same (research or clinical) purposes. Importantly, the individual items of the AVHRS-Q also corresponded highly to the items of the AVHRS, with the exception of four moderately correlating items. Given that the AVHRS-Q had to be short and not (too) cognitively demanding, explanations, and examples of items were not included in the self-report questionnaire. This may have resulted in discrepancies (and therefore

moderate correlations) between four specific items of the AVHRS-Q and AVHRS. It is therefore important to keep in mind that whilst the AVHRS-Q can be used to reliably achieve a quick overall severity measure of AVH (similar to the interview-based AVHRS), one should be cautious when only interpreting single items of the AVHRS-Q (specifically the items on form of address, severity of negative content, interference with daily functioning, interference with thoughts). Moreover, the AVHRS-Q severity scores were only moderately related to measures of quality of life and psychological distress, which indicates that the AVHRS-Q specifically measures characteristics and severity of AVH. Overall, the AVHRS-Q demonstrates good convergent and divergent validity in this study.

An important feature of the AVHRS-Q is that it takes only 6 min on average to complete. This makes it exceptionally suitable for quick and frequent assessments, for instance in research on the effectiveness of treatments or for frequent monitoring in a clinical context, such as routine outcome monitoring (ROM) assessments. Currently, ROM assessments for patients with psychosis often consist of more global outcome measures for positive symptoms, such as the positive and negative syndrome scale (PANSS) [32–34]. One PANSS item assesses the severity of hallucinations, but does not inquire about, for example, whether the patient has separate or simultaneous voices, whether the patient has negative or positive voices, or even how AVH interfere with daily functioning. All these aspects may be potentially relevant for treatment or in signifying the nature of distress. Currently, the AVHRS-Q is being utilized in the ROM protocol of the Voices Outpatient Department of the University Medical Center Groningen.

The current study has some limitations. First, in contrast to the HPSVQ [17] and the DV-SA [18], the AVHRS-Q does not enquire about the social circumstances of AVH, or whether the command hallucinations are obeyed. However, the AVHRS-Q does enquire about the interference with daily functioning and the presence of command hallucinations, which can be further explored during therapy. Second, similar to the validation study on the AVHRS interview [10] we did not measure sensitivity to change. As all patients were in therapy for their voices and the AVHRS-Q is incorporated in treatment, retest data would likely be confounded with therapeutic effects. To assess this in an unbiased manner, a control group not receiving treatment for their AVH should have been included. However, given that all patients had quite severe AVH for a substantial amount of years, this was deemed unethical. Third, the current study recruited two reasonably chronic patient samples, implying the current findings may be less generalizable to healthier populations. However, the AVHRS-Q has already been administered in a general population sample, supporting its use in less chronic samples [25].

One important strength of the current study is that the AVHRS-Q is based on an existing measure, the AVHRS, which has already been deemed to have good psychometric properties [10] and was used in multiple research projects [26, 30]. A second strength is that the AVHRS-Q was evaluated by patients with AVH and that their feedback was used to improve the AVHRS-Q into its current form. Third, given that AVH are prevalent in multiple disorders, therefore, being a trans-diagnostic symptom, it is a strength that the AVHRS-Q was tested in patients with different disorders.

To conclude, the AVHRS-Q is a quick self-report version of a validated interview on auditory hallucinations already in use, the AVHRS. The current study demonstrates that the AVHRS-Q has good internal consistency, convergent validity and divergent validity. The AVHRS-Q can very well be applied in both clinical practice and research, where it is required to assess AVH in a quick and reliable manner.

**Acknowledgements** This research received no specific grant from any funding agency, commercial or non-profit sectors. The authors are most grateful to all participants in this study. The authors thank Rosemarie de Boer MSc, Chris Geraets MSc, Ingrid Tewelde-Kampert MSc and Erna van 't Hag MSc for informing their patients about the study, and master students from the University of Groningen for their help with the interviews and retrieval of diagnoses. We thank Professor Edwin van den Heuvel for calculating the sample size for Sample I.

## Compliance with ethical standards

**Conflict of interest** On behalf of all authors, the corresponding author states that there is no conflict of interest.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Kelleher I, Connor D, Clarke MC et al (2012) Prevalence of psychotic symptoms in childhood and adolescence: a systematic review and meta-analysis of population-based studies. *Psychol Med* 42:1857–1863. <https://doi.org/10.1017/S0033291711002960>
- Linscott RJ, van Os J (2013) An updated and conservative systematic review and meta-analysis of epidemiological evidence on psychotic experiences in children and adults: on the pathway from proneness to persistence to dimensional expression across mental disorders. *Psychol Med* 43:1133–1149. <https://doi.org/10.1017/S0033291712001626>
- de Leede-Smith S, Barkus E (2013) A comprehensive review of auditory verbal hallucinations: lifetime prevalence, correlates and mechanisms in healthy and clinical individuals. *Front Hum Neurosci* 7:367. <https://doi.org/10.3389/fnhum.2013.00367>
- Bartels-Velthuis AA, van de Willige G, Jenner JA et al (2011) Course of auditory vocal hallucinations in childhood: 5-year follow-up study. *Br J Psychiatry* 199:296–302. <https://doi.org/10.1192/bjp.bp.110.086918>
- Larøi F, Sommer IE, Blom JD et al (2012) The characteristic features of auditory verbal hallucinations in clinical and nonclinical groups: state-of-the-art overview and future directions. *Schizophr Bull* 38:724–733. <https://doi.org/10.1093/schbul/sbs061>
- Frederick JA, Killeen MR (1998) Instruments for assessment of auditory hallucinations. *Arch Psychiatr Nurs* 12:255–263. [https://doi.org/10.1016/S0883-9417\(98\)80035-6](https://doi.org/10.1016/S0883-9417(98)80035-6)
- Bartels-Velthuis AA, Jenner JA, van de Willige G et al (2010) Prevalence and correlates of auditory vocal hallucinations in middle childhood. *Br J Psychiatry* 196:41–46. <https://doi.org/10.1192/bjp.bp.109.065953>
- Kelleher I, Cederlöf M, Lichtenstein P (2014) Psychotic experiences as a predictor of the natural course of suicidal ideation: a Swedish cohort study. *World Psychiatry* 13:184–188. <https://doi.org/10.1002/wps.20131>
- Spencer C, Castle D, Michie PT (2002) Motivations that maintain substance use among individuals with psychotic disorders. *Schizophr Bull* 28:233–247. <https://doi.org/10.1093/oxfordjournals.schbul.a006934>
- Bartels-Velthuis AA, van de Willige G, Jenner JA, Wiersma D (2012) Consistency and reliability of the auditory vocal hallucination rating scale (AVHRS). *Epidemiol Psychiatr Sci* 21:305–310. <https://doi.org/10.1017/S2045796012000108>
- Chadwick P, Lees S, Birchwood M (2000) The revised beliefs about voices questionnaire (BAVQ-R). *Br J Psychiatry* 177:229–232. <https://doi.org/10.1192/bjp.177.3.229>
- Haddock G, McCarron J, Tarrier N, Faragher EB (1999) Scales to measure dimensions of hallucinations and delusions: the psychotic symptom rating scales (PSYRATS). *Psychol Med* 29:879–889. <https://doi.org/10.1017/S0033291799008661>
- Ratcliff K, Farhall J, Shawyer F (2011) Auditory hallucinations: a review of assessment tools. *Clin Psychol Psychother* 18:524–534. <https://doi.org/10.1002/cpp.729>
- Birchwood M, Meaden A, Trower P et al (2000) The power and omnipotence of voices: subordination and entrapment by voices and significant others. *Psychol Med* 30:337–344. <https://doi.org/10.1017/S0033291799001828>
- Mann B, Pakenham KI (2006) Development of a measure to assess coping for auditory hallucinations. *Aust J Psychol* 58:93–100. <https://doi.org/10.1080/00049530600730450>
- Chadwick P, Barnbrook E, Newman-Taylor K (2007) Responding mindfully to distressing voices: links with meaning, affect and relationship with voice. *Spec Issue Towar New Underst Psychosis* 44:581–587
- Van Lieshout RJ, Goldberg JO (2007) Quantifying self-reports of auditory verbal hallucinations in persons with psychosis. *Can J Behav Sci* 39:73–77. <https://doi.org/10.1037/cjbs2007006>
- Pinto A, Gigantesco A, Morosini P, La Pia S (2007) Development, reliability and validity of a self-administered questionnaire on subjective opinion about delusions and voices. *Psychopathology* 40:312–320. <https://doi.org/10.1159/000105529>
- Jenner J, van de Willige G (2002) The auditory vocal hallucination rating scale (AVHRS). University Medical Center Groningen, Groningen
- de Jesus DR, Gil A, Barbosa L et al (2011) A pilot double-blind sham-controlled trial of repetitive transcranial magnetic stimulation for patients with refractory schizophrenia treated with clozapine. *Psychiatry Res* 188:203–207. <https://doi.org/10.1016/j.psychres.2010.11.022>
- Lambert MJ, Burlingame GM, Umphress V et al (1996) The reliability and validity of the outcome questionnaire. *Clin Psychol Psychother* 3:249–258. [https://doi.org/10.1002/\(SICI\)1099-0879\(199612\)3:4%3C249::AID-CPP106%3E3.0.CO;2-S](https://doi.org/10.1002/(SICI)1099-0879(199612)3:4%3C249::AID-CPP106%3E3.0.CO;2-S)

22. Derogatis LR, Rickels K, Rock AF (1976) The SCL-90 and the MMPI: a step in the validation of a new self-report scale. *Br J Psychiatry* 128:280–289. <https://doi.org/10.1192/bjp.128.3.280>
23. Priebe S, Huxley P, Knight S, Evans S (1999) Application and results of the manchester short assessment of quality of life (MANSA). *Int J Soc Psychiatry* 45:7–12. <https://doi.org/10.1177/002076409904500102>
24. Van de Willige G, Bartels-Velthuis AA, Jenner JA (2010) The Auditory Vocal Hallucination Rating Scale Questionnaire (AVHRS-Q). University of Groningen, Groningen
25. Bartels-Velthuis AA, Wigman JTW, Jenner JA et al (2016) Course of auditory vocal hallucinations in childhood: 11-year follow-up study. *Acta Psychiatr Scand*. <https://doi.org/10.1111/acps.12571>
26. Bartels-Velthuis AA, van de Willige G, Jenner JA et al (2012) Auditory hallucinations in childhood: associations with adversity and delusional ideation. *Psychol Med* 42:583–593. <https://doi.org/10.1017/S0033291711001590>
27. IBM (2014) IBM SPSS statistics for windows, version 23.0. IBM Corp, Armonk
28. Cronbach LJ (1951) Coefficient alpha and the internal structure of tests. *Psychometrika* 16:297–334. <https://doi.org/10.1007/BF02310555>
29. Steenhuis LA, Bartels-Velthuis AA, Jenner JA et al (2016) Religiosity in young adolescents with auditory vocal hallucinations. *Psychiatry Res* 236:158–164
30. Bartels-Velthuis AA, Blijd-Hoogewys EMA, van Os J (2011) Better theory-of-mind skills in children hearing voices mitigate the risk of secondary delusion formation. *Acta Psychiatr Scand* 124:193–197. <https://doi.org/10.1111/j.1600-0447.2011.01699.x>
31. Maijer K, Palmén SJMC, Sommer IEC (2017) Children seeking help for auditory verbal hallucinations; who are they? *Schizophr Res* 183:31–35. <https://doi.org/10.1016/j.schres.2016.10.033>
32. Egger ST, Vetter S, Weniger G et al (2016) The use of the health of the nation outcome scales for assessing functional change in treatment outcome monitoring of patients with chronic schizophrenia. *Front Public Health* 4:1–6. <https://doi.org/10.3389/fpubh.2016.00220>
33. Tasma M, Swart M, Wolters G et al (2016) Do routine outcome monitoring results translate to clinical practice? A cross-sectional study in patients with a psychotic disorder. *BMC Psychiatry* 16:107–114. <https://doi.org/10.1186/s12888-016-0817-6>
34. Kay SR, Fiszbein A, Opler LA (1987) The positive and negative syndrome scale for schizophrenia. *Schizophr Bull* 13:261–276. <https://doi.org/10.1093/schbul/13.2.261>