# University of Groningen

## Social predictors of psychotic experiences in adolescence

Steenhuis, Laura Alida

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

# The Development, Validity and Reliability of the Auditory Vocal Hallucination Rating Scale Questionnaire (AVHRS-Q)
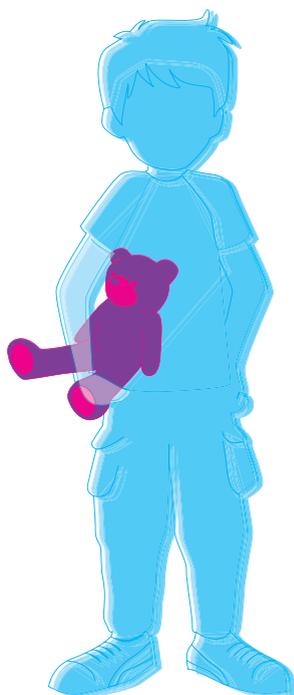
Steenhuis, L. A., Pijnenborg, G. H. M., van der Willige, G., Visser, E., Nauta, M. H., Aleman, A. & Bartels-Velthuis, A. A. (2019)

## 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 was 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 = .90, p<0.01). The AVHRS-Q and AVHRS did not differ in the identification of mild and severe voice-hearers (X2 (1, N = 32) = 15.71). AVHRS-Q severity scores had moderate correlations with measures of psychological distress (OQ-45, r = .43, p<0.01; SCL-90, r = .50, p<0.05) and quality of life (MANSA, r = -.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.

## 2.1 Introduction

Auditory vocal hallucinations (AVH) are prevalent in children, adolescents and adults, both in clinical settings and in the general population (Bartels-Velthuis, et al., 2011; de Leede-Smith & Barkus, 2013; Kelleher et al., 2012; Linscott & van Os, 2013). AVH are common in psychotic illnesses and other mental disorders such as depression, bipolar -, dissociative -, and substance use disorders (Larøi et al., 2012). AVH severity may be predictive of (amongst others) social problems (Bartels-Velthuis et al., 2010; Frederick & Killeen, 1998), suicidal ideation (Kelleher et al., 2014) or substance abuse (Spencer, Castle, & Michie, 2002) and a reliable assessment of AVH is therefore very important. 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 by 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 (Bartels-Velthuis, van de Willige, Jenner, & Wiersma, 2012) 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; Chadwick, Lees, & Birchwood, 2000) and the number of voices (in contrast to the PSYRATS; Haddock, McCarron, Tarrier, & Faragher, 1999) (see the validation paper by Bartels-Velthuis, van de Willige, Jenner, & Wiersma, 2012 for a more elaborate description). Given that there has been a shift from interview measures to self-report measures of AVH (Ratcliff et al., 2011), 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 (Ratcliff et al., 2011). These questionnaires are usually tailored to measure a specific aspect of AVH, for example beliefs about AVH (e.g. BAVQ-R; Chadwick et al., 2000), interpretations and attitudes towards AVH (e.g. VPD; Birchwood, Meaden, Trower, Gilbert, & Plaistow, 2000), coping with AVH (RAHQ; Mann & Pakenham, 2006), and mindfulness of AVH (SMVQ; Paul Chadwick, Barnbrook, & Newman-Taylor, 2007). 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); Van Lieshout & Goldberg, 2007; 10 items, The Delusion and Voices Self-Assessment (DV-SA); Pinto, Gigantesco, Morosini, & La Pia, 2007). 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; Bartels-Velthuis, van de Willige, Jenner, & Wiersma, 2012; Jenner & van de Willige, 2002), 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; Bartels-Velthuis, van de Willige, Jenner, & Wiersma, 2012), demonstrating good convergent validity. As greater severity of AVH is related to both increased psychological distress and a lower quality of life (de Jesus et al., 2011), it is expected that the severity measure of the AVHRS-Q will correlate with measures of psychological distress (the Outcome Questionnaire, OQ-45; Lambert et al., 1996 and the Symptom Checklist, SCL-90; (Derogatis, Rickels, & Rock, 1976) and quality of life (the Manchester Short Assessment of Quality of Life, MANSA; Priebe, Huxley, Knight, & Evans, 1999). 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.

## 2.2 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 18-65 years old, and having a good command of the 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) in order 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 was 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 this data was 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 was 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.

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

|  | Sample I (n=32) | Sample II (n=82) |
|---|---|---|
| Female (n, %) | 18 (56) | 44 (54) |
| Age (*M*, SD) | 38 (11) | 39 (12) |
| Duration of AVH in years (*M*, SD) | 12 (11) | N.A. |
| Psychotic disorder[$] (n, %) | 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) |

*Note.*

*AVH; Auditory Vocal Hallucinations. N.A.; Not Available.*

[$] *Psychotic disorders in these samples consisted of Schizophrenia, Schizoaffective disorder and Psychotic disorder not otherwise specified.*

**Measures**

**The AVHRS and Development of the AVHRS-Q.** The AVHRS (Bartels-Velthuis, van de Willige, Jenner, & Wiersma, 2012; Jenner & van de Willige, 2002) is a structured 16-item interview, administered by an experienced therapist to evaluate AVH during a period of one month. The AVH are rated on 4- and 5-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 (Van de Willige et al., 2010) 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 http://www.rgoc.nl/#home/downloads (see table 2 for a summary of the items). The AVHRS-Q has 17 items, 15 of which are assessed with a 4- and 5-point scale and two on a 10-point scale. For the 4- and 5- point scales, scores range from 0 (not applicable) to 3 or 4 (most applicable). For the 10-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, one hour, and several hours to continuously) in comparison to five options in the questionnaire (see table 2, item 4). Moreover, the wording of some items were 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 (Bartels-Velthuis, et al., 2011; Bartels-Velthuis et al., 2016; Bartels-Velthuis, van de Willige, Jenner, Wiersma, et al., 2012; Bartels-Velthuis et al., 2010), 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; (Bartels-Velthuis, van de Willige, Jenner, & Wiersma, 2012). 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).

*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 minutes | 0 |
| | 15 minutes 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 (1$^{st}$ person) | 0 |
| | The voices speak to me (2$^{nd}$ person) | 0 |
| | The voices speak to me and about me (2$^{nd}$ and 3$^{rd}$ person) | 1 |
| | The voices speak about me (3$^{rd}$ 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 |
| | Much louder than own voice (yelling, shouting, screaming) | 1 |
| 9. Positive or Negative Content | Always positive | 0 |
| | Mostly positive but occasionally negative | 0 |
| | Mostly neutral or equally positive and negative | 0 |
| | Mostly negative | 1 |
| | Always negative | 1 |

| | | |
|---|---|---|
| 10. Severity of Negative Content | 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 |
| | Most of the time | 1 |
| | All of the time | 1 |
| 13. Interference with thoughts | Never | 0 |
| | Occasionally | 0 |
| | Half of the time | 1 |
| | Most of the time | 1 |
| | All of the time | 1 |

| 14. Control | Always | 0 |
|---|---|---|
| | Most of the time | 0 |
| | Half of the time | 0 |
| | Less than half of the time | 1 |
| | Never | 1 |
| 15. Attribution of Origin | The voices only concern myself | 0 |
| | The voices mostly concern myself | 0 |
| | Equally concern myself as outside influences | 1 |
| | Mostly caused by outside influences | 1 |
| | Fully caused by outside influences | 1 |
| 16. Frequency of distress (1= never, 10=always) | 0-6 | 0 |
| | 7-10 | 1 |
| 17. Intensity of distress (1= not at all, 10=extremely) | 0-6 | 0 |
| | 7-10 | 1 |
| Total: | | Minimum=0 Maximum=14 |

*Note.*
*AVHRS-Q; Auditory Vocal Hallucination Rating Scale-Questionnaire.*
*This table represents a summary of the questionnaire and is not suitable for assessments.*
*The full questionnaire can be downloaded at: http://www.rgoc.nl/#home/downloads.*

**Quality of Life.** Quality of life was assessed with the Manchester Short Assessment of Quality of Life (MANSA; Priebe et al., 1999, a 16-item self-report measure consisting of four objective items and 12 subjective items (satisfaction with life, accommodation, 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 7-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 was 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 (Lambert et al., 1996) 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 5-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 (Derogatis et al., 1976) is a 90-item self-report measure, assessing a variety of psychopathology. Each item is rated on a 5-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 (IBM, 2014). 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).

In order 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 (Cronbach, 1951).

In order 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.

## 2.3 Results

### Descriptives
In Sample I, the AVHRS-Q took an average of 5.8 minutes to be completed (SD: 2.72, range: 2-15), whereas the AVHRS took an average of 14.3 minutes 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 measure of the AVHRS and AVHRS-Q were highly correlated. This correlation did not differ for participants who started with the AVHRS (r = .90, p<0.01) and those who started with the AVHRS-Q (r = .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 .72 (see table 4). The mean AVH severity measure and the distribution of severity groups did nog differ significantly between the AVHRS-Q and the AVHRS.

*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 $r$ | Paired Samples t-test | Exact Mc Nemar's test |
|---|---|---|---|---|---|---|
| AVH Severity score (*M, SD*) |  | 6.91 (3.15) | 7.38 (3.66) | $r$=.90** | $t$(31)=-.55 |  |
| AVH Severity Group (*N, %*) | Mild | 24 (75%) | 22 (68.8%) |  |  | $X^2$ (1, N = 32) = 15.71 |
|  | Severe | 8 (25%) | 10 (31.3%) |  |  |  |

Note:
*p<0.05, **p<0.01
AVH; Auditory Vocal Hallucinations, AVHRS-Q; Auditory Vocal Hallucination Rating Scale-Questionnaire, AVHRS; Auditory Vocal Hallucination Rating Scale.

*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 | .86** |
| 2.   Separately or simultaneously | .84** |
| 3.   Frequency | .73** |
| 4.   Duration | .70** |
| 5.   Hypnagogic and/or hypnopompic voices[$] | |
| 6.   Location | .71** |
| 7.   Form of address | .56** |
| 8.   Loudness | .78** |
| 9.   Positive or Negative Content | .85** |
| 10. Severity of Negative Content | .57** |
| 11. Anxiety | .74** |
| 12. Interference with Daily Functioning | .46** |
| 13. Interference with thoughts | .52** |
| 14. Control | .82** |
| 15. Attribution of Origin | .84** |
| 16. Frequency of distress | .44* |
| 17. Intensity of distress | .70** |

*Note.*
*\*p<0.05, \*\*p<0.01*
*AVHRS-Q; Auditory Vocal Hallucination Rating Scale-Questionnaire, AVHRS; Auditory Vocal Hallucination Scale, $could not be computed as all participants consistently reported option 4 'at all times of the day'.*

**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) = .46, p>0.05) and SCL-90 (t(80) = -.48, p>0.05)

*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)*

| | | MANSA | SCL-90 | OQ-45 |
|---|---|---|---|---|
| | *M (SD)* | (n=82) | (n=24) | (n=62) |
| AVH Severity score (Pearson's *r)* | 7.66 (2.69) | -0.22** | .50* | .43** |

Note.
*\*p<0.05, \*\*p<0.01*
*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.*

## 2.4 Discussion

The current study shows that the Auditory Vocal Hallucination Rating Scale Questionnaire (AVHRS-Q) (Van de Willige et al., 2010) 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; Bartels-Velthuis, van de Willige, Jenner, & Wiersma, 2012). In addition, the AVHRS-Q is shown to be a specific measure of AVH and not a general measure of psychological distress (OQ-45; Lambert et al., 1996) and SCL-90; Derogatis et al., 1976) or quality of life (MANSA; Priebe et al., 1999). 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 (Bartels-Velthuis, Blijd-Hoogewys, & van Os, 2011; Bartels-Velthuis, van de Willige, Jenner, & Wiersma, 2012; Maijer, Palmen, & Sommer, 2017; Steenhuis et al., 2016) 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 six minutes 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) (Egger et al., 2016; Kay, Fiszbein, & Opler, 1987; Tasma et al., 2016). 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 (Van Lieshout & Goldberg, 2007) and the DV-SA (Pinto et al., 2007), 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 (Bartels-Velthuis, van de Willige, Jenner, & Wiersma, 2012) 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. In order 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 (Bartels-Velthuis et al., 2016).

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 (Bartels-Velthuis, van de Willige, Jenner, & Wiersma, 2012) and was used in multiple research projects (Bartels-Velthuis, et al., 2011; Bartels-Velthuis, van de Willige, Jenner, Wiersma, et al., 2012). 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.

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.