Chapter 3

Course of auditory vocal hallucinations in childhood: a 5-year follow-up study

Agna A. Bartels-Velthuis
Gerard van de Willige
Jack A. Jenner
Jim van Os
Durk Wiersma

British Journal of Psychiatry (accepted for publication)
Abstract

Background
In a baseline study among 7-8 year old children with auditory vocal hallucinations (AVH), only limited functional impact was observed.

Aims
To assess AVH 5-year course and its predictors, as well as AVH 5-year incidence and its risk factors.

Method
A sample of 337 12- and 13-year-old children were reassessed on AVH and associated symptoms after a mean follow-up period of 5.1 years.

Results
The AVH 5-year persistence and incidence rates were 24% and 9% respectively, more new cases arising in urban areas. Both persistent and incident AVH were associated with problem behaviour in the (CBCL) clinical range, particularly at follow-up, as well as with other psychotic symptoms, particularly at baseline. AVH persistence was predicted by baseline AVH severity, particularly in terms of external attribution of voices and hearing multiple voices, and was associated with worse primary school test scores and lower level secondary school.

Conclusions
First-onset AVH in middle childhood is neither rare nor neutral in terms of psychopathological and behavioural comorbidity. Persistence of AVH in childhood similarly is not rare and associated with psychopathological, behavioural and cognitive alterations.

Declaration of interest
None.
Introduction

Subclinical psychotic experiences in the general population are prevalent in both children and adults, and not generally associated with persistence over time or onset of psychiatric disorder. Nevertheless, literature reviews suggest that a small number will make the transition to a clinical psychotic disorder (Kaymaz and Van Os, 2010a; Van Os et al., 2009).

In a previous study on a representative sample of 7-8 year old children in the Netherlands, the prevalence of auditory vocal hallucinations (AVH) was 9% (Bartels-Velthuis et al., 2010). Although AVH in these children mostly were of limited functional impact, a subgroup with serious suffering was considered at risk for more severe psychotic outcomes.

The current study presents a 5-year follow-up of this sample (now 12/13 years of age) and examines the persistence and new incidence of AVH, as well as their clinical relevance in terms of problem behaviour, AVH severity and associations with aetiological variables. In view of the baseline results and previous work in this area, it was hypothesized that (i) AVH severity at baseline would predict persistence (Askenazy et al., 2007; Escher et al., 2002b; Poulton et al., 2000), (ii) AVH would show associations with behavioural problems (Eminson, 2007; Janssens et al., 2010; Welham et al., 2009b) and other psychotic symptoms (Ott et al., 2001), (iii) AVH incidence and severity would be associated with environmental factors like cannabis use (Henquet et al., 2005; Hides et al., 2009; Mackie et al., 2010) and urbanicity (Cougnard et al., 2007; Spauwen et al., 2006a; Weiser et al., 2007), (iv) baseline AVH characteristics suggesting external attribution and higher level of intrusion would be predictive of persistence (Askenazy et al., 2007; Escher et al., 2002b; Poulton et al., 2000) and (v) AVH would be associated negatively with cognitive ability (Cannon et al., 2002; Horwood et al., 2008a; Jabben et al., 2007; Welham et al., 2009a).

Method

Procedure

From the case-control sample of the first wave (n = 694, of which 347 with AVH) parents of 605 children (87%; of which 50% with AVH) gave informed consent for follow-up. These parents were sent a notification letter by mail. Non-responders were sent a reminder followed by a second letter, if necessary. In case of persisting non-response, parents were contacted by telephone if their numbers
could be traced. Seven female interviewers (six bachelor students from the department of Orthopedagogy of the University of Groningen and one psychology graduate) received extensive training by the first three authors in conducting the interviews. First, they were introduced to the topic of auditory hallucinations, and then, with consent, observed several therapeutic sessions of patients receiving treatment at the Voices Outpatient Department (VOPD) of the University Medical Center Groningen. Interviewers were informed in detail about the structure and results of the baseline study and received training in the administration of instruments and in conducting interviews with children. Finally, they interviewed patients of the VOPD, under supervision of the second and third author. In addition to the formal interview training, several booster sessions were arranged to discuss interview and scoring procedures and to prevent interviewer ‘drift’. Interviewers were instructed in a detailed protocol on how to approach families and how to conduct the interviews. In order to prevent bias, interviewers were unaware of children’s AVH status at baseline. Children were interviewed at home, separately from their parents. First, children were screened on the experience of hearing voices in the past five years; AVH-positive children were subsequently interviewed with the Auditory Vocal Hallucination Rating Scale (AVHRS) (Jenner and Van de Willige, 2002). Given evidence for associations between auditory hallucinations and use of substances (Arseneault et al., 2004; Henquet et al., 2005) and alterations in performance at school (Cannon et al., 2002), information was also gathered in these domains. Parents completed the Child Behaviour Check List, designed for children 4-18 years of age (CBCL; Achenbach, 1991; Dutch translation Verhulst et al., 1995), and a socio-demographic questionnaire. Finally, parents were instructed on how to contact the research team in case they or their children had questions or worries resulting from the interview.

**Subjects**
A total of 337 children were interviewed at follow-up, representing 49% of the baseline sample of 694 and 56% of the 605 that had consented to follow-up. Parents of 48 children were willing to complete the CBCL without their children being interviewed, 112 parents withdrew their informed consent and 108 families could not be contacted because of non-response or (r)emigration. Children at follow-up, similar to baseline, continued to be evenly distributed over baseline case or control status: 170 AVH-positive children and 167 control children. From the combined baseline and follow-up data, four groups could be
distinguished: (i) children hearing voices at baseline and during the five-year follow-up period: the persistent group, (ii) children hearing voices at baseline but not during the five-year follow-up period: the remitted group, (iii) children not hearing voices at baseline but positive for AVH over the follow-up period: the incident group, and (iv) children not hearing voices both at baseline and at follow-up: the referent group. The persistent AVH group was defined as: children hearing voices in the year before the baseline interview and (at least occasionally) during one other year over the follow-up period. The incident group was defined as children who heard voices for the first time after baseline assessment, with a minimum duration of three months.

_Instruments_

_Children_
The 16-item AVHRS is a structured interview (Jenner and Van de Willige, 2002), rating on a 5-point scale characteristics of hearing voices (e.g., frequency, attribution, duration, loudness, negative content, distress, anxiety, control, interference with thinking and with daily life). Psychometric properties of the AVHRS are good (Bartels-Velthuis et al., 2008). During booster sessions, DVD-recorded AVHRS interviews (of consenting patients attending the VOPD) were rated by all interviewers and the first author. Total inter-rater agreement score (weighted Cohen’s kappa) was 0.88.

Use of substances (cannabis, synthetic drugs like XTC, amphetamines, cocaine and heroin) was assessed by a self-report questionnaire, in order to avoid socially desirable answers.

_Cognitive ability_

In order to rate children’s performance at school, children and parents were asked to provide the results of the national Dutch ‘end-of-primary-school test’, which assesses the continuation level for secondary education. The standard scores range from 1–50; children scoring 1–28 are qualified for vocational secondary education, and those scoring 29–50 for pre-academic secondary education. In addition, current secondary school level was assessed.

getParents

The CBCL/4-18 (Achenbach, 1991) is a self-report questionnaire for parents regarding their child’s behaviour. It consists of 113 items, grouped into nine syndrome scales of behavioural, social and physical functioning. Items are rated on a three-point scale (‘0’ = ‘not true’; ‘1’ = ‘somewhat or sometimes true’; ‘2’ =
‘very true or often true’). Parents also provided demographic data including their educational level and family income.

**Analysis**

Conform baseline analysis (Bartels-Velthuis et al., 2010), an AVHRS severity index was computed by recoding items to ‘0’ = ‘none or mild consequences’ versus ‘1’ = ‘considerable to severe consequences’. Based on this index, two groups were defined: a ‘severe AVH’ group (children scoring ≥ 5) and a ‘mild AVH’ group (scores 0–4), that were contrasted with the referent group.

Degree of urbanization was defined by the family’s home address and split (conform baseline analysis [Bartels-Velthuis et al., 2010]) in rural (somewhat urban and rural: ‘0’) and urban (very, strongly and moderately urban: ‘1’), based on data of the official classification of urbanicity provided by Statistics Netherlands (2009a).

From the CBCL total score, the borderline clinical range of psychopathology was determined according to Achenbach (1991), i.e. a cut-off score of 31 for both sexes.

In order to avoid multiple testing, only CBCL main indices (internalizing, externalizing, total score and clinical range) were used. In addition, the ‘thought problems’ scale was used, as the items of this scale represent psychotic experiences and behaviour (‘Strange behaviour’, ‘Strange ideas’, ‘Hears sounds or voices that aren’t there’, ‘Sees things that aren’t there’, ‘Stares blankly’, ‘Can’t get his/her mind off certain thoughts’, ‘Repeats certain acts over and over’). Given that the item ‘hearing things’ overlaps with the dependent variable of AVH, analyses with the ‘thought problems’ scale were also carried out without this item, as well as with a more narrow selection representing the three (non-AVH) psychotic experiences: ‘Strange behaviour’, ‘Strange ideas’ and ‘Sees things that aren’t there’.

Secondary education was defined at two levels: low (‘1’ = ‘vocational’) and high (‘0’ = ‘pre-academic’). Socio-economic status (SES) was derived from parental averaged educational levels and family income, resulting in a three-level variable of low (= ‘1’), middle (= ‘2’) and high (= ‘3’) SES.

Analyses were carried out using SPSS for Windows, version 16.0. Standard multinomial logistic regression, yielding odds ratios (OR) and 95% confidence intervals (CI), were used to compare the four groups of incident, persistent, remitted and referent children, with the referent group as reference category. Significance tests were two-tailed with alpha set at .05.

Differential non-response due to differences in help-seeking and psychiatric
service use was tested by tracing the children anonymously at a group level through the Psychiatric Case Register North-Netherlands (PCR-NN) over the study period.

Results

**Sample characteristics and AVH incidence and prevalence rates**

Mean age of the follow-up sample was 13.1 years (SD = 0.5, range 12.0–14.6). Mean interval between baseline and follow-up assessment was 5.1 years (SD = 0.4). Attrition analyses (based on n = 694) showed that the participation rate was equal for baseline AVH-positive children (49%) and controls (48%). Registration rates at the PCR-NN were not different for participating and non-participating children (8% vs. 10%; OR = 0.83, 95% CI 0.49–1.40, P = .48), neither did these groups differ in mean number of psychiatric service contacts: 2.5 (SD = 1.3) in the participating children and 2.4 (SD = 1.6) in the non-participating children (OR = 1.06, 95% CI 0.75–1.51, P = .73). The participation rate for girls was significantly higher than for boys (53% vs. 44%; χ² = 5.50, P = .02). There was no evidence for differential attrition as a function of age and urbanicity. A flow chart of follow-up participants (n = 337) is presented in figure 1.

Most children in the follow-up sample (90%) lived in a rural environment. Socio-economic status (SES) was evenly distributed (31% low, 39% middle and 30% high). Mean score on ‘end-of-primary-school test’ was 37.2 (SD = 8.7, range 9–50); 53% of the children were attending higher level secondary education. In Table 1, demographic data by AVH follow-up status are presented.

**Course of AVH and predictors of persistence**

The AVH persistence rate was 23.5% and the 5-year cumulative incidence rate in the baseline control sample was 9.0%. Mean duration of hearing voices in the persistent group was 5.7 years (SD = 1.9), in the incident group this was 3.5 years (SD = 1.7).

Children still hearing voices after five years, compared to children in the remitted group, at baseline more often belonged to the group with severe AVH (40% vs. 21%, OR = 2.54, 95% CI 1.19–5.45, P = .016) (Table 2). Examining distinct baseline AVH characteristics revealed that hearing more than one voice discriminated best between persistence and remission of baseline AVH (OR = 2.81, 95% CI 1.09–7.23, P = .032). Also attribution of voices to an external source (e.g., deceased family members or an extra-terrestrial source) discriminated between persistence and remission (OR = 2.35, 95% CI 1.11–4.97, P = .025).
**Figure 1.** Flow chart of AVH in baseline and follow up participants

**AVH and cognitive ability**
Children with persistent AVH, compared to the referent group, displayed lower end-of-school cognitive test scores (mean score 32 vs. 37; OR = 1.07, 95% CI 1.02–1.12, P = .004) and more often attended a lower level secondary school (71% vs. 48%; OR = 2.63, 95% CI 1.21–5.68, P = .014). The incident group, however, did not differ from the referent group. Children in the remitted group scored highest on both cognitive ability outcome measures (Table 2).

**Problem behaviour**
Compared to the referent group, children with persistent AVH, as well as children with incident AVH, displayed higher scores on the ‘thought problems’ scale of the CBCL at both baseline and at follow-up; Table 3). Also when the item on auditory hallucinations (item 40) was excluded from this scale these differences remained (‘thought problems’ baseline: persistent AVH: OR = 1.53, 95% CI 1.07–2.19, P = .02; incident AVH: OR = 1.82, 95% CI 1.10–3.02, P = .02; ‘thought problems’ follow-up: persistent AVH: OR = 1.27, 95% CI 0.93–1.75, P = .14; incident AVH: OR = 1.82, 95% CI 1.29–2.55, P = .001). Odds ratio’s were highest when only the narrow psychotic symptoms of the ‘thought problems’ scale (i.e. ‘strange behaviour’, ‘strange ideas’, and ‘sees things that aren’t there’) were included: baseline: persistent AVH: OR = 2.40, 95% CI 1.15–5.02, P = .020; incident AVH: OR = 3.19, 95% CI 1.33–7.62, P = .009; follow-up: persistent AVH: OR = 2.37, 95% CI 1.10–5.10, P = .027; incident AVH: OR = 4.06, 95% CI 1.84–8.97, P = .001). CBCL
scores at follow-up were more often in the clinical range for both the persistent and the incident AVH group (Table 3).

**AVH severity**

At baseline, 27% of the AVH-positive children \( (n = 347) \) reported severe suffering associated with auditory hallucinations, against 35% of AVH-positive children at follow-up \( (n = 55) \). The proportion of children with severe AVH at follow-up was higher in the incident group (60%) than in the persistent group (25%) \( (\chi^2 = 5.91, P = .02) \). Children with severe AVH (but not children with mild AVH), compared to the referent group, were more likely to show problem behaviour (CBCL total score: baseline: \( OR = 1.03, 95\% CI 1.00–1.05, P = .033; \) follow-up: \( OR = 1.04, 95\% CI 1.02–1.06, P = .000) \). At follow-up, children with severe AVH were more likely to score in the clinical range of psychopathology \( (OR = 5.45, 95\% CI 2.10–14.16, P = .000) \). AVH severity at follow-up was not associated with gender \( (\chi^2 = 1.09, P = .58) \), SES \( (\chi^2 = 3.49, P = .48) \) or urbanicity \( (\chi^2 = 4.23, P = .38) \). Severe AVH children had lower end-of-primary-school test scores than mild AVH children; both groups had lower scores than control children \( (F = 6.69, P = .001) \). Severe children (and not mild) were more likely to attend a lower secondary school type than the referent group \( (OR = 4.40, 95\% CI 1.41–13.73, P = .01) \).

**Risk factors**

Compared to the referent group, children with incident AVH more often lived in an urban environment \( (OR = 3.86, 95\% CI 1.08–13.89, P = .038) \). Severity of AVH at follow-up was not associated with urbanicity \( (OR 1.02, 95\% CI 0.23–4.63, P = .98) \). Substance use in the sample was negligible; those who ever used cannabis (the only substance used) were in the incident \( (n = 1) \) and referent \( (n = 2) \) groups.
### Table 1. Sample characteristics at follow up (n = 337)

<table>
<thead>
<tr>
<th></th>
<th>Incident (n = 15)</th>
<th>Persistent (n = 40)</th>
<th>Remitted (n = 130)</th>
<th>Referent (n = 152)</th>
<th>test</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years) mean</td>
<td>13.2</td>
<td>13.1</td>
<td>13.1</td>
<td>13.2</td>
<td>F = 0.59</td>
<td>n.s.</td>
</tr>
<tr>
<td>Gender (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>female</td>
<td>60</td>
<td>55</td>
<td>53</td>
<td>54</td>
<td>χ² = 0.28</td>
<td>n.s.</td>
</tr>
<tr>
<td>Socio-economic status (SES) (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>low</td>
<td>47</td>
<td>35</td>
<td>32</td>
<td>27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>middle</td>
<td>27</td>
<td>40</td>
<td>35</td>
<td>44</td>
<td>χ² = 5.06</td>
<td>n.s.</td>
</tr>
<tr>
<td>high</td>
<td>27</td>
<td>25</td>
<td>33</td>
<td>29</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2. Multinomial regression analyses of cognitive abilities and AVH severity (n = 337)

<table>
<thead>
<tr>
<th>Cognitive ability</th>
<th>Incident (n = 15)</th>
<th>Persistent (n = 40)</th>
<th>Remitted (n = 130)</th>
<th>Referent* (n = 152)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% mean (SD)</td>
<td>OR (CI)</td>
<td>% mean (SD)</td>
<td>OR (CI)</td>
</tr>
<tr>
<td>end-of-primary school test score*</td>
<td>35.33 (6.43)</td>
<td>1.03 (0.96-1.10)</td>
<td>31.90 (8.93)</td>
<td>1.07 (1.02-1.12)</td>
</tr>
<tr>
<td>secondary school, low level</td>
<td>47 % (0.32-2.71)</td>
<td>.90 (.90)</td>
<td>71 % (2.63)</td>
<td>.14 (1.21-5.68)</td>
</tr>
<tr>
<td>Severity AVH</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Severe AVH (baseline)**</td>
<td>-</td>
<td>40 % (2.54)</td>
<td>21 % (1.19-5.45)</td>
<td>.016</td>
</tr>
<tr>
<td>Severe AVH (follow-up)*****</td>
<td>60 % (4.50)</td>
<td>.019 (1.28-15.81)</td>
<td>25 % (4.50)</td>
<td>.016</td>
</tr>
</tbody>
</table>

Abbreviations: AVH, Auditory Vocal Hallucinations; SD, Standard Deviation; OR, Odds Ratio; CI, Confidence Interval.
* the reference comparison group in the multinomial regression analysis.
** the reference group is remitted AVH.
*** the reference group is persistent AVH.
* OR's were inversed (1/OR) so that higher OR's reflect worse performance.
### Table 3. Multinomial regression analyses of CBCL indices ($n = 337$)

<table>
<thead>
<tr>
<th></th>
<th>Incident ($n = 15$)</th>
<th>Persistent ($n = 40$)</th>
<th>Remitted ($n = 130$)</th>
<th>Referent* ($n = 152$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% mean (SD) OR (CI)</td>
<td>% mean (SD) OR (CI)</td>
<td>% mean (SD) OR (CI)</td>
<td>% mean (SD) OR (CI)</td>
</tr>
<tr>
<td><strong>Problem behaviour (baseline CBCL)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thought problems</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- without item on AVH</td>
<td>1.8 (2.3) 1.92 (1.25-2.97) .003</td>
<td>1.3 (1.7) 1.66 (1.21-2.27) .002</td>
<td>0.6 (1.1) 1.13 (0.86-1.48) .40</td>
<td>0.5 (0.8)</td>
</tr>
<tr>
<td>- only 3 psychotic symptoms</td>
<td>1.4 (2.3) 1.82 (1.10-3.02) .02</td>
<td>1.0 (1.3) 1.53 (1.07-2.19) .02</td>
<td>0.4 (0.8) 0.91 (0.65-1.28) .59</td>
<td>0.5 (0.8)</td>
</tr>
<tr>
<td>Internalizing</td>
<td>0.7 (1.5) 3.19 (1.33-7.62) .009</td>
<td>0.4 (0.9) 2.40 (1.15-5.02) .02</td>
<td>0.1 (0.3) 0.98 (0.43-2.27) .97</td>
<td>0.1 (0.3)</td>
</tr>
<tr>
<td>Externalizing</td>
<td>7.5 (7.4) 1.06 (0.95-1.18) .31</td>
<td>7.2 (7.6) 1.05 (0.98-1.12) .14</td>
<td>5.7 (6.2) 1.01 (0.98-1.06) .72</td>
<td>5.6 (4.8)</td>
</tr>
<tr>
<td>Total</td>
<td>5.3 (5.3) 0.94 (0.82-1.09) .43</td>
<td>9.7 (7.3) 1.05 (1.00-1.10) .07</td>
<td>8.4 (7.9) 1.03 (0.99-1.07) .17</td>
<td>7.1 (6.5)</td>
</tr>
<tr>
<td>Clinical range</td>
<td>23.8 (21.8) 1.01 (0.97-1.05) .64</td>
<td>29.9 (21.3) 1.03 (1.01-1.05) .02</td>
<td>22.8 (19.4) 1.01 (0.99-1.02) .42</td>
<td>21.0 (14.8)</td>
</tr>
<tr>
<td><strong>Problem behaviour (follow-up CBCL)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thought problems</td>
<td>1.9 (2.6) 1.68 (1.26-2.25) .050</td>
<td>1.0 (1.3) 1.33 (1.02-1.72) .03</td>
<td>0.5 (1.1) 0.96 (0.76-1.22) .75</td>
<td>0.5 (1.0)</td>
</tr>
<tr>
<td>- without item on AVH</td>
<td>1.6 (2.1) 1.82 (1.29-2.55) .001</td>
<td>0.7 (1.1) 1.27 (0.93-1.75) .14</td>
<td>0.4 (1.0) 0.99 (0.76-1.29) .93</td>
<td>0.4 (0.8)</td>
</tr>
<tr>
<td>- only 3 psychotic symptoms</td>
<td>0.7 (1.3) 4.06 (1.84-8.97) .001</td>
<td>0.2 (0.6) 2.37 (1.10-5.10) .03</td>
<td>0.1 (0.4) 1.29 (0.61-2.69) .51</td>
<td>0.1 (0.3)</td>
</tr>
<tr>
<td>Internalizing</td>
<td>11.9 (10.2) 1.15 (1.07-1.23) .000</td>
<td>9.3 (7.1) 1.11 (1.05-1.17) .000</td>
<td>5.7 (5.5) 1.01 (0.97-1.06) .64</td>
<td>5.4 (4.7)</td>
</tr>
<tr>
<td>Externalizing</td>
<td>8.7 (13.6) 1.05 (0.99-1.12) .11</td>
<td>7.1 (6.1) 1.03 (0.98-1.08) .25</td>
<td>5.2 (5.4) 0.98 (0.94-1.03) .43</td>
<td>5.7 (5.9)</td>
</tr>
<tr>
<td>Total</td>
<td>35.5 (33.4) 1.04 (1.02-1.06) .001</td>
<td>28.3 (19.7) 1.03 (1.01-1.04) .006</td>
<td>18.9 (15.9) 1.00 (0.98-1.01) .79</td>
<td>19.4 (14.8)</td>
</tr>
<tr>
<td>Clinical range</td>
<td>47.4 (46.7) 1.55 (14.08) .006</td>
<td>33.3 (25.7) 2.57 (1.16-5.67) .020</td>
<td>19.0 (16.3) 2.12 (0.65-2.25) .55</td>
<td>16.0 (14.8)</td>
</tr>
</tbody>
</table>

Abbreviations: CBCL, Child Behaviour Checklist; AVH, Auditory Vocal Hallucinations; SD, Standard Deviation; OR, Odds Ratio; CI, Confidence Interval.
Discussion

Summary of findings
The majority (76%) of the children with auditory vocal hallucinations at age 7 or 8 years no longer heard voices at the age of 12 and 13 years. This is in line with remission rates presented in a recent systematic review (Van Os et al., 2009), which reported that approximately 75–90% of developmental psychotic experiences are transitory and disappear over time. Although the samples in the systematic review all involved older children (from around 11 years of age), the current study suggests that the natural transitory course also applies to children aged 7 to 8 years at baseline. The current results suggest that AVH may influence variables indexing cognitive ability, as these variables showed strong associations with persistence, but not with incidence. Persistence was associated with greater severity of AVH at baseline with some attenuation by the time of follow-up, when incident AVH were associated with the greatest level of severity. In line with this was the finding that the incident AVH group displayed a nearly five times, and the persistent AVH group an approximately 2.5 times greater risk for CBCL scores in the clinical range at follow-up. Psychotic thought problems were associated with incident and persistent AVH, both at baseline and at follow-up. However, the higher OR’s for psychotic ideation in the incidence group suggest a higher risk in this group (Dominguez et al., 2009; Smeets et al., 2010). These differences between the incident and the persistent AVH groups suggests that (i) persistent AVH may improve over time (the proportion in the severe range dropped from 40% at baseline to 25% at follow-up) and (ii) the first manifestation of AVH in later childhood may be indicative of more severe underlying pathology.

Findings and initial hypotheses
The course of AVH did not always follow an orderly pattern. Thus, children in the persistent group had been hearing voices for an average period of more than 5.7 years, and all had heard voices in the past year. Approximately 50% of the persistent group heard voices on a more or less regular (continuous) basis, some of them as long as they could remember. However, in other children, the course seemed intermittent during the 5-year follow-up period. Despite the underlying variability in frequency and duration, persistence of hearing voices more often occurred in children with more severe AVH at baseline, confirming the first hypothesis.

The second hypothesis was also confirmed. The pattern of results was that
baseline psychotic symptoms were associated with incidence and persistence of AVH, and had strong associations with CBCL scores in the clinical range. These findings suggest that AVH that are associated with a broader range of psychotic symptoms fare worse and give rise to a clinical problem behaviour. At baseline, the rate of AVH was higher in rural areas; however, AVH were of greater severity and had more functional impact in the urban area. At follow-up, children in the incident group more often lived in an urban environment, but AVH severity was not associated with an urban environment, thus only partly confirming the third hypothesis. At baseline, level of urbanicity was established by primary school postal code because children’s home addresses were not available. This procedure seemed justified because primary school addresses mostly were in close range with the children’s home address. At follow-up however, home addresses were available, which we now used as estimates because the majority of children living in a rural area would now attend secondary schools in an urban environment.

Logistic regression analyses with baseline AVH characteristics as covariates showed that hearing several voices (more than one) and the attribution of these voices to an external cause were the strongest predictors of hearing voices after five years. This confirms our fourth hypothesis, in line with findings of Escher and colleagues (2002b), suggesting that children with a secondary attribution of their voices (i.e., the child had indicated explanations of the voices being caused by spirits, ghosts, special gifts etc.) had a higher risk of voice persistence. Previous work suggests that individuals who continue to hear voices and attribute AVH to an external source may be more at risk of developing secondary psychotic ideation (Krabbendam et al., 2004b).

Cognitive ability (indexed by end-of-primary-school test and secondary school level) of children with persistent AVH was significantly lower compared to children whose AVH were only transitory, and thus may be indicative of risk of transition to psychotic disorder, as suggested by recent studies (Cannon et al., 2002; Dominguez et al., 2010; Horwood et al., 2008a; Jabben et al., 2007; Van Oel et al., 2002; Weiser et al., 2009; Welham et al., 2009a). The incident group did not (yet) show alterations in cognitive ability. However, they did have a lower mean score on the end-of-primary-school test (albeit non-significantly), which arguably represents a better measure of cognitive ability than secondary school level, given that many other factors intervene in the choice for secondary school type. In addition, the number of children in the incident group was small (n = 15). Taken together, these findings provide some support for the fifth hypothesis.

At baseline, associations were tested using prospective data on pre- and perinatal
complications, derived from routine Infant Health Service records at baseline assessment (Bartels-Velthuis et al., 2010); no clear pattern of association was found. Variables tested at baseline included dichotomous measures on pregnancy, delivery and condition of the child right after birth, as well as continuous variables on early developmental characteristics (in the child’s first twelve months of life), summarised in three variables for analysis: fine motor activity, gross motor activity and communication (Bartels-Velthuis et al., 2010). Post-hoc examination of these perinatal and developmental variables in relation to AVH groups at follow-up similarly revealed no significant association for any of the variables (data not shown) indicating that the absence of a clear pattern of association between AVH and these variables at baseline was valid.

**Clinical implications**
Based on the above results, it may be concluded that the majority of auditory hallucinations at baseline are benign and will be transient. If presented in a context of behavioural problems or with other psychotic experiences, it may be advisable for parents to seek help for their child, in order to prevent a poor prognosis. However, parental recognition of the child’s auditory hallucinations was low. At baseline, 13% of the parents indicated that their child may hear ‘things’, whereas at follow-up, only 25% of the parents of the persisting voice hearing children indicated correct recognition of the phenomenon. For clinicians, the findings may be helpful in the sense that in case of children presenting with somatic complaints, behavioural problems, lower school results or thought problems, the possibility of underlying auditory hallucinations should be considered. On the other hand, when a child reveals the existence of uncomplicated auditory hallucinations, explanations of the mostly transitory nature of the experience may be appropriate.

**Limitations**
We could not include all children of the baseline case-control sample. First, 13% of the parents had already indicated at baseline that they did not want to participate in a follow-up study. Second, despite maximum efforts to include as many children as possible, this was not always possible as parents were sometimes protective towards their children who had just started secondary school and were in a process of adjustment. Third, there is good reason to assume that a number of children did not want to participate a priori. Fourth, some of the non-responding families could not be traced or contacted by telephone. There were, however, no indications that non-response or refusal rates were due to
(psychiatric) problems of the children. Participation rates of baseline AVH-positive children and controls were almost equal (50%). Likewise, the follow-up participation rate of baseline severe and mild AVH-positive children equals the proportion at baseline (25% severe and 75% mild). These considerations were confirmed by data of the Psychiatric Case Register North-Netherlands, in which referral and contact rates at group level also showed no differences. Finally, the study design did not permit us to gather any clinical information, including diagnostic assessments by treatment providers. This would have shed more light on the association between childhood auditory hallucinations and clinical disorders.