Hearing Voices in a Non-Psychiatric Population

Catherine Lawrence

Solihull Child and Adolescent Mental Health Service, UK

Jason Jones

Thames Valley Forensic Mental Health Service and University of Birmingham, UK

Myra Cooper

University of Oxford, UK

Background: Many people hear voices but do not access psychiatric services and their experiences are largely unknown, not least because of the difficulty in contacting such people. This study investigates the beliefs held about voices, distress experienced, and provides a topographical account of the experience of hearing voices in a sample of individuals who hear voices in a non-psychiatric population. Method: A quantitative questionnaire internet-based study with a within-subjects and between-subjects design was used. The internet was used to make contact with people who hear voices. One hundred and eighty-four participants in the general population who heard voices completed measures online assessing anxiety, depression, and beliefs about their dominant voice. Results: Participants reported a broad range of experiences associated with hearing voices, some in keeping with the previous research on clinical populations. Conclusion: The use of the internet to recruit and research non-clinical samples of people who hear voices is supported. This study provides details regarding demographic information and the experience of voice hearing from a fairly large sample of people who hear voices in a non-psychiatric population. It lends support to the idea that voice hearing occurs on a continuum, with evidence that many people hear voices in the general population and are not distressed by the experience.

Keywords: Hearing voices, auditory hallucination, general population, schizophrenia.

Introduction

The experience of hearing voices has traditionally been associated with symptoms of mental health disorders. Strauss (1969) proposed a continuum of psychosis that promotes a normalizing approach to people’s experiences, thereby reducing the stigma associated with
distinguishing people as “different” from the general population solely on the basis of a perceptual experience (Johns and van Os, 2001, p. 1137). For this to carry any merit it relies on the presence and experience of hearing voices within a non-clinical population.

Romme and Escher (1989) were the first to recognize the experience of hearing voices as a phenomenon in its own right and not simply as a form of symptomatology. Romme, Honig, Noorthoorn and Escher (1992) investigated the experience of hearing voices in the general population. Their results seemed to indicate that the judgments individuals made about their voices were salient in their experience of distress. Individuals unable to cope with their voices were significantly more likely to be in psychiatric care compared to those who reported being able to cope (Romme et al., 1992). The notion of an individual coping with voice hearing promoted a departure from the medical model.

Cognitive models of voice hearing and distress (e.g. Chadwick, Birchwood and Trower, 1996) have begun to explain how and why individuals might become distressed about their experience of hearing voices and offer suggestions of how to intervene clinically to reduce this distress. Chadwick and Birchwood (1994) suggest that the beliefs about a voice, rather than voice content, determine whether a voice is perceived as malevolent or benevolent, and thus whether the voice hearer is distressed and the voice is resisted or engaged with.

To date, most studies on voice hearing within the general population have been constrained by the examination of the symptom rather than the experience (e.g. Romme et al. 1992; Verdoux et al., 1998). This potentially limits the exposition of the experience, with questions taking a narrow, symptomalogical focus.

Studies that have investigated the distribution of hearing voices in the general population have used different measures to determine whether someone is hearing voices and have reported their findings differently. General population surveys report the annual prevalence of hearing voices at between 0.6 and 1.2 percent (Bak et al., 2005; Johns, Nazroo, Bebbington and Kuipers, 2002 respectively), while the lifetime prevalence has been found to be 4.3% (Eaton, Romanoski, Anthony and Nestadt, 1991). Figures from specific non-clinical populations and student samples produce estimates between 4.8% and 71% of participants hearing voices (Verdoux et al., 1998; Posey and Losch, 1983 respectively). Overall, however, these findings support the existence of voice hearing in non-clinical groups and lends potential support to the conceptualization of voice hearing as being on a continuum (Strauss, 1969).

Some studies used psychiatric measures and diagnostic criteria to determine whether people were hearing voices. This may have excluded people who heard voices whose experiences would not be classed as an auditory hallucination according to nosological systems, so the figures for prevalence may be artificially low in some studies. A general population study that was not restricted to psychiatric diagnostic categories or specific samples may provide more generalizable findings.

Studying hallucinatory experiences in the general population is likely to inform the understanding of the experience in the clinical population. Allen et al. (2005) investigated hallucinatory predisposition in a sample of 327 students using an internet-based study. Given some estimates of low prevalence rates of the experience of hearing voices, this study established a useful method of capturing data on a large sample.

Conducting research online is still in the early stages but is becoming more popular, with 1,133,408,294 people connected to the internet worldwide (Miniwatts Marketing Group, 2007). Rodham and Gavin (2006) identified the internet as an important method of recruitment for participants from difficult-to-reach populations, such as those to be studied here.
A limitation of using the internet is that there is no fail-safe way of ensuring that the data collected have come from the respondent (Rodham and Gavin, 2006). However, this is not a problem exclusive to researchers using the internet as all researchers are dependent on participants providing honest, reliable answers, even when completing questionnaires or interviews in person (Rodham and Gavin, 2006).

Rodham and Gavin (2006) suggest that the internet may be well-suited to collect data as participants may “feel freer to express their ‘true’ feelings” due to the anonymity it provides (p. 94). As well as having a disinhibiting effect, anonymity also reduces social desirability and thus encourages “true” and “authentic” responses from participants (Rodham and Gavin, 2006, p. 95). The internal validity of internet-based research may also be improved due to the elimination of interviewer effects (Mustanski, 2001). Rodham and Gavin (2006) conclude that conducting research online “poses no more ethical dilemmas than when conducting research by more traditional means” (p. 96).

The present study aims to investigate the experience of hearing voices in a non-psychiatric sample. To maximize recruitment of participants, an internet-based method of data collection was used. The primary focus of the investigation was to generate an account of the experience of hearing voices that could be readily compared to recent research on the characteristics of voices and associated distress in clinical samples. It was hypothesized that distress among a general population sample would be lower than that found in clinical populations; that endorsement of beliefs about the omnipotence and malevolence of the voice would be less than previously published data for clinical populations; and that participants would endorse higher levels of benevolence of related beliefs than clinical populations.

**Method**

**Apparatus**

The website that collected the data was written using PHP version 4.3 and questions and responses were stored in a MySQL 4.0.18 database.

**Participants**

Three hundred and ninety-six people took part in the study online. Of these, 184 met the inclusion criteria of being adults who currently heard voices, or had heard voices in the past. Exclusion criteria were individuals who were under the age of 18 years, were seeking or had sought psychiatric help for their voices, heard voices when under the influence of substances, or could not read English.

**Procedure**

Individuals interested in taking part in the study accessed the website at their own convenience. The homepage contained the advert for the study and a link to the participant information page. Participants consented to take part by selecting the “yes” option to state that they agreed to the informed consent statement. Participants were then taken through six pages of questions/questionnaires and two pages requiring demographic information. After submitting their responses they were taken to a thank you and debriefing page.
Completing the study took an average of 12.59 minutes (range of 5–75 minutes), with the modal value being 10 minutes.

**Measures**

Participants were screened regarding their experiences of hearing unexplained noises or voices using five questions. They were asked to indicate if they currently did, had in the past, or had ever had the experiences when not drinking alcohol or taking drugs. Only participants who endorsed at least one of the following three items as “currently I do” or “I have in the past” were considered to be individuals who experienced hearing voices. The definition of hearing voices used was: hearing faint or clear sounds of people or a person mumbling or talking when there is no one around; or, hearing a voice speak when there was no one around; or hearing lots of people talking when there was no one about. These questions were developed from those used in previous studies (e.g. Posey and Losch, 1983; Johns et al., 2002). Questions did not take a psychiatric perspective in order to widen the experience to the general population. They also focused on hearing voices rather than thoughts or other unexplained noises to promote focus on the experience of hearing voices.

Questionnaires were presented, one to a page, with option buttons for participants to mark their responses. Forced choice responding, with all questions requiring a response before the participant could move to the next page, ensured that the problem of missing data was eliminated. Only one response per question was allowed.

**Emotional distress: anxiety and depression**

The Hospital Anxiety and Depression Scale (HADS) (Zigmond and Snaith, 1983) is a standardized assessment of recent emotional distress. It was designed for use with general medical outpatients but is reported to be valid in community-based samples. The HADS was chosen over other standardized measures of anxiety and depression (e.g. Beck scales) as it was designed for use with a non-psychiatric population, has fewer items for people to complete so there is more chance they will continue with the study, has been used in conjunction with the measure of beliefs about voices also used in this study (Chadwick, Lees and Birchwood, 2000), has published psychometric properties for internet administration (Andersson, Kåldö-Sandström, Ström and Strömgren, 2003; McCue, Buchanan and Martin, 2006) and the use of it online has been found to have “comparable . . . psychometric properties” to paper-and-pencil administration (Andersson et al., 2003, p. 259).

**Voices: beliefs and demographics**

The revised Beliefs about Voices Questionnaire (BAVQ-R) (Chadwick et al., 2000) measures beliefs, feelings and behaviour associated with the experience of hearing voices in five subscales (malevolence, benevolence, omnipotence, resistance, engagement). The full BAVQ-R was completed. The voices demographics section asked participants when they had last heard their voice, how many voices they heard, the gender of their main voice, and whether they knew the identity of their voice. It also included three questions from the Topography of Voices Rating Scale (TVRS) (Hustig and Hafner, 1990), those items that assess frequency, volume and clarity of voices.
Participant demographics
Participants were asked to provide some basic demographic details. Gender and age were the only compulsory questions.

Results

Gender and age
The participants consisted of 126 females (68.5%) and 58 males (31.5%). Their mean age was 34.52 years ($SD = 13.08$), with a mode of 18 years and range of 18 to 66 years.

Ethnicity
The majority of participants considered themselves White (84.8%). Other people described their ethnicity as Black African or Black Caribbean (1.6%), Dual heritage (4.9%), or Asian (3.8%). Some participants chose not to disclose ethnicity (4.9%).

Relationship status
The majority of participants were single (34.8%); the next most common relationship status was married (27.7%), living with partner (16.8%), and not living with partner (10.3%), divorced (6.5%), and separated (1.6%). Four participants did not indicate their relationship status.

Employment status
The majority of participants were in full-time employment (40.8%). The next most common status of employment was to be a full-time student (15.8%), and be in part-time paid employment (12.5%). Fifteen participants were unemployed (8.2%) and nine were receiving sickness/incapacity/invalidity benefits (4.9%). Four participants did not provide their employment status. Participants had jobs in administration, business, childcare, cleaning, computing, education, healthcare, information, management, and sales.

Country
Eighty-four participants (45.7%) were living in the United Kingdom and 54 (29.3%) in the United States. The remaining participants were from other European, Asian, American and Australasian countries.

Method of recruitment
Participants were recruited mainly via the internet (64.7%) through Google searches, adverts on websites, and a link on a voice supporter’s discussion group. Forty-one participants were recruited via e-mail (22.3%). One participant was recruited via a poster, 20 participants selected “other”, and three did not state how they heard about the study.
Table 1. Scores on the HADS (N = 184).

<table>
<thead>
<tr>
<th></th>
<th>Anxiety</th>
<th>Depression</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% (N)</td>
<td>% (N)</td>
</tr>
<tr>
<td>Normal range 0-7</td>
<td>37.0 (68)</td>
<td>57.1 (105)</td>
</tr>
<tr>
<td>Mild range 8-10</td>
<td>16.3 (30)</td>
<td>24.5 (45)</td>
</tr>
<tr>
<td>Moderate range 11-14</td>
<td>20.0 (36)</td>
<td>10.9 (20)</td>
</tr>
<tr>
<td>Severe range 15-21</td>
<td>27.2 (50)</td>
<td>7.6 (14)</td>
</tr>
</tbody>
</table>

*Note: HADS = Hospital Anxiety and Depression Scale.*

**Last heard voice**

The majority of participants had last heard a voice within 24 hours of completing the study (38.0%) or within the previous week (15.2%); 53.2 percent had therefore heard voices within the past week. Twelve percent of participants last heard a voice over 4 years ago and 4.3 percent stated that they were unsure.

**Number of voices**

Fifty-one percent of participants heard only one voice, 14% heard two different voices, 16.2% heard between three and eight voices, and 5.4% heard over 10 voices. Twenty-four participants (13.0%) stated that they were unsure.

**Distress**

The mean HADS anxiety subscale score was 10.24 (SD = 5.45). The anxiety subscale was internally consistent, yielding a Cronbach’s alpha coefficient of .90. The mean HADS depression subscale score was 6.39 (SD = 4.88). The depression subscale was also internally consistent, with a Cronbach’s alpha of .87. The scores from participants expressed in clinical ranges on the HADS are depicted in Table 1.

Pooled variance t-tests were conducted to determine if there were statistically significant differences between mean scores on subtests for the participants and a clinical sample of individuals who heard voices. Scores for anxiety and depression were significantly lower for the participants compared to Chadwick et al.’s (2000) clinical sample of people who heard voices (t = 3.40, df = 240, p < .001; t = 5.44, df = 240, p < .001 respectively). Nevertheless, it is worthy of note that the mean score on the anxiety subscale lies at the top end of the “mild” range according to the normative data available for the HADS.

**Beliefs about voices**

Table 2 gives the descriptive statistics¹ and Cronbach’s alphas for the BAVQ-R subscales. All scales had good internal reliability. Participants used the full range of scores for all subscales.

---

¹The medians and modes are shown as not all subscales followed a normal distribution.
on the BAVQ-R. Compared to a clinical sample of 71 people who heard voices (Chadwick et al., 2000), the sample in this study had significantly lower mean scores on subscales of malevolence \( t = 8.37, df = 253, p < .001 \), omnipotence \( t = 7.46, df = 253, p < .001 \) and resistance \( t = 8.33, df = 253, p < .001 \), and significantly higher mean scores on subscales of benevolence \( t = -3.31, df = 253, p = .001 \) and engagement \( t = -2.06, df = 253, p < .05 \). Some of the subscales were not normally distributed but there was no alternative statistical test to the Pooled variance \( t \)-test so results should be interpreted with caution.

A strong positive relationship between omnipotence and malevolence was found \( r = .73, p < .01 \), and between omnipotence and resistance \( r = .59, p < .01 \). No significant relationship was found between omnipotence and engagement \( r = .07, p = .17 \) nor between omnipotence and benevolence \( r = .12, p = .09 \). A strong positive relationship was found between malevolence and resistance \( r = .75, \) one-tailed, \( p < .01 \) and benevolence and engagement \( r = .86, \) one-tailed, \( p < .01 \). Negative significant relationships were found between malevolence and benevolence \( r = -.25, \) one-tailed, \( p < .01 \), engagement and resistance \( r = -.43, \) one-tailed, \( p < .01 \), malevolence and engagement \( r = -.23, \) one-tailed, \( p < .01 \) and benevolence and resistance \( r = -.41, \) one-tailed, \( p < .01 \).

Positive associations were identified between levels of distress and malevolence, omnipotence and resistance.

### Gender and identity of voice

Sixty-five participants (35.3\%) stated that their most dominant voice was male, 48 (26.1\%) said that it was female and 71 (38.6\%) were unsure. There was no significant relationship between the gender of the individual and the perceived gender of the voice. There was no statistically significant difference in level of anxiety or depression based on the gender of the voice (Anxiety: \( \chi^2 = 5.30, df = 2, p = .07 \); Depression: \( \chi^2 = 3.22, df = 2, p = .20 \)). Seventy percent of participants did not know the identity of their most dominant voice. However, people who did not know the identity of their dominant voice reported higher levels of anxiety \( U = 2870.00, p < .05 \) and depression \( U = 2610.50, p < .01 \). Further, participants rated higher levels of malevolence \( U = 2893.50, p < .05 \) and resistance \( U = 2827.50, p < .05 \) for voices whose identity was unknown, whereas benevolence \( U = 2318.00, p < .001 \) and engagement \( U = 2602.00, p < .01 \) were rated higher if the identity was known.
Topography of voices

Twenty participants (10.9%) heard their voice every hour, 25.5% several times a day, 6.5% once a day, 20.1% several times a week, and 37% had not heard it lately. For 67 participants (36.4%), their voices were “normal” in loudness, 8.7% experienced them as “very loud” and 16.8% as “fairly loud”. Fifty-five participants (29.9%) felt their voices were “fairly quiet” and 8.2% “very quiet”. Seventy-four participants (40.2%) experienced their voices as “very clear” and 21.2% as “fairly clear”, while 9.2% thought they were “very mumbled” and 15.2% “fairly mumbled”, with 14.1% experiencing the clarity of their voice as “normal”.

People who reported hearing their voice more frequently reported higher levels of anxiety ($\chi^2 = 25.90, df = 4, p < .001$), depression ($\chi^2 = 27.94, df = 4, p < .001$), malevolence ($\chi^2 = 42.41, df = 4, p < .001$), omnipotence ($\chi^2 = 56.41, df = 4, p < .001$), and resistance ($\chi^2 = 33.97, df = 4, p < .001$).

Discussion

The majority of participants fell within the normal range for anxiety and depression on the HADS. Scores for anxiety and depression were also significantly lower compared to a clinical sample of people who heard voices. Thus, people who hear voices in the general population appear to be less distressed than those in receipt of psychiatric services. However, like the relevant clinical studies, the present study has not focused on what the individual is distressing themselves about. There is no way of telling whether the distress reported in studies is specifically associated with the experience of hearing voices. Evidently, however, the experience of voices per se is not associated with clinical levels of distress. Thus we would caution about the assumption that the experience of hearing voices is inherently pathological or indicative of disorder.

The topographical information found in this sample permits interesting comparisons to clinical population studies. In addition to the information about the perceptual experience of the voice (clarity, volume), the frequency and recency of voice hearing suggests that the experience is often one that is continuous, rather than a one-off experience. This supports the notion that the perceptual nature of voice hearing in a non-psychiatric population is similar to that of clinical populations. Thus, again we would assert that it is not the experience of voices per se that is associated with pathology. The most notable difference was the gender of the voice. In the present study, similar proportions were found for the gender of the voice, which is at odds with the clinical population studied by Junginger and Frame (1985). They reported that the majority of voices were perceived as male. Further, there was no evident relationship between the gender of the individual and the perceived gender of their dominant voice.

The majority of participants in the present study did not know the identity of their voice. Inferences about malevolence and omnipotence were more likely if the individual did not know the identity of their voice, whereas inferred benevolence and subsequent engagement of the voice was more likely if the identity was known. This might suggest that participants in a non-clinical sample perceive their voices to be more kindly and are more prepared to engage with them than individuals in a clinical sample. This has implications for further research about the relationship between the individual and their voice. In cognitive behavioural interventions for voice hearing, it is this relationship that is often important to examine. The
inferred or announced identity of the voice may be closely associated with inferences about the malevolence or omnipotence of the voice.

Anonymous responding was used to ensure maximum participation but this did not allow for validation and reliability checks among participants responding. Therefore, people who falsely claim to hear voices may have completed the questionnaires. This is a potentially critical flaw of the methodology. As the study did not contain any “lie-detection” items it can not be certain that people completed the questionnaires truthfully.

In an endeavour to increase the sample diversity and reduce the possibility of bias, recruitment via several methods was attempted as recommended by Mustanski (2001). However, despite the use of posters, leaflets and adverts, most people were recruited via the internet and e-mail. Although the internet allowed a wider sample of people to be reached, the internet still only reaches 17.2% of the world population (Miniwatts Marketing Group, 2007). This may affect the generalizability of the results, as the demographics of those with such access may be very different from those without, for example, in terms of education, employment and culture.

In order to complete the study participants first had to have access to a computer and use of the internet; they also had to be literate, or have someone help them complete the questionnaires. Participants who accessed the study via Google searches were actively searching the internet for information on hearing voices. As such, this sample is self-selecting, which may be biased towards those particularly interested in the topic. However, evaluating self-selection bias is difficult with this population as there is limited information about the “non-participants” to compare them to the participants (Costigan and Cox, 2001, p. 707).

Conducting research online yielded more participants in a fairly short space of time than was expected. Recruitment proved more successful via adverts on search engines compared to internet groups, posters and leaflets. Whilst more time consuming in the early phases of creating the website, data collection, scoring and analysis was made easier with people completing questionnaires in their own time and data being scored within the created database. For a difficult to reach population such as this, recruitment and participation via the internet has proved to be a successful method.

The questionnaires used were designed for clinical populations and asked participants to consider their responses in view of their dominant voice. If people in the general population do not experience a single voice as dominant, this may make it harder for them to answer the questions. This is a limitation of all voice hearing research that uses measures that ask participants to focus on their dominant voice (Birchwood et al., 2004). All of the questionnaires used in this study were created and normed in the UK but this study was accessible to the English-speaking world. Cultural differences may have affected subscale scores. In order to compare the responses of the present sample to a clinical population, the present data were compared to previously published data from Chadwick et al. (2000). As the samples were contacted in different ways, had different histories, and had had contact with mental health services (in the Chadwick et al., 2000, study). We recognise that the comparison made is subject to error. Future research should make such comparisons under stricter controls.

This study also did not assess the content of voices, largely because much previous research (e.g. Birchwood and Chadwick, 1997) has suggested that content is not solely important in the distress experienced by people who hear voices.

This study highlighted a number of areas for future research. Although the present study went some way to exploring the experience of hearing voices in a non-psychiatric population,
further work could be undertaken to explore the meaning of the experience. This might include an analysis of how many people attribute their experience to religion, spirituality or some other phenomenon. Research addressing the actual emotional responses to the perceptual inference that is hearing voices would be useful in the non-psychiatric population, as this might confirm hypothesized relationships between healthy and unhealthy responses to the experience of hearing voices.

To date, this is the largest study of the beliefs held about voices in a non-psychiatric population. The use of the internet to approach and research people who hear voices in the general population has been supported. This study demonstrates that adults across the lifespan, in varying types of employment and relationships, throughout the world, experience hearing voices when there is no one around and do not access psychiatric services. It provides support to the idea that voice hearing occurs on a continuum, with evidence that many people hear voices in the general population and are not distressed by the experience.

Acknowledgments

Thanks to all the participants, to Neil Lawrence for his computing skills, and to Peter Hopkinson from the Learning Disability Speciality, Northamptonshire Healthcare NHS Trust, for the time to write this paper.

References


