The Beliefs about Voices Questionnaire – Revised: A factor structure from 450 participants

Clara Strauss\textsuperscript{a,b,\*}, Kenneth Hugdahl\textsuperscript{c,d,e}, Flavie Waters\textsuperscript{f,g}, Mark Hayward\textsuperscript{h,h}, Josef J. Bless\textsuperscript{c,e}, Liv E. Falkenberg\textsuperscript{c,e}, Bodil Kråkvik\textsuperscript{h}, Arve Egil Asbjørnsen\textsuperscript{c}, Erik Johnsen\textsuperscript{d,e,i}, Igne Sinkeviciute\textsuperscript{d,e,j}, Rune A. Krokend\textsuperscript{e,i}, Else-Marie Løberg\textsuperscript{d,e,i,k,l}, Neil Thomas\textsuperscript{m,n}

\textsuperscript{a} School of Psychology, University of Sussex, Brighton BN1 9QH, UK
\textsuperscript{b} Sussex Partnership NHS Foundation Trust, Hove, UK
\textsuperscript{c} Department of Biological and Medical Psychology, University of Bergen, Norway
\textsuperscript{d} Division of Psychiatry, Haukeland University Hospital, Bergen, Norway
\textsuperscript{e} NORMENT Center of Excellence, University of Oslo, Norway
\textsuperscript{f} Clinical Research Centre, North Metro Health Service, Grayslands Hospital, Perth, Australia
\textsuperscript{g} School of Psychiatry and Clinical Neurosciences, University of Western Australia, Australia
\textsuperscript{h} Nidaros District Psychiatric Center, Department of Research and Development, St. Olavs University Hospital, Trondheim, Norway
\textsuperscript{i} Department of Addiction Medicine, Haukeland University Hospital, Bergen, Norway
\textsuperscript{j} Centre for Research and Education in Forensic Psychiatry, Haukeland University Hospital, Bergen, Norway
\textsuperscript{k} Department of Addiction Medicine, Haukeland University Hospital, Bergen, Norway
\textsuperscript{l} Department of Clinical Psychology, University of Bergen, Norway
\textsuperscript{m} Centre for Mental Health, Swinburne University of Technology, Melbourne, Australia
\textsuperscript{n} Voices Clinic, Monash Alfred Psychiatry Research Centre, Alfred Hospital and Monash University Central Clinical School, Melbourne, Australia

\textbf{ARTICLE INFO}

Keywords:
Hallucinations
Hearing voices
BAVQ
Beliefs about voices
Omnipotence
Malevolence

\textbf{ABSTRACT}

Hallucinated voices are common across psychiatric and non-clinical groups. The predominant cognitive theory about the impact of voices posits that beliefs about voice power (‘Omnipotence’) and voice intent (‘Malevolence’/‘Benevolence’) play a key role in determining emotional and behavioral reactions. The revised Beliefs about Voices Questionnaire (BAVQ-R) was designed to assess these constructs, together with two styles of responding (Engagement and Resistance). The BAVQ-R is widely used in clinical and research settings, yet it has not received validation of its constructs and factor structure. This study examined the factor structure of the BAVQ-R by combining datasets from five study centers, comprising 450 participants (belief constructs) and 269 participants (response styles), and using confirmatory and exploratory factor analysis. Findings failed to support a three factor belief model, instead showing a two-factor structure (‘Persecutory beliefs’ combining Omnipotence and Malevolence constructs, and a Benevolent construct). Emotional and behavioral items did not separate. Overall, results showed that (i) a two-factor model of beliefs (Persecutory and Benevolent beliefs) provides a better fit to the data than a three-factor model, and (ii) emotional and behavioral modes of responding items should not be separated. Theoretical implications of this finding are discussed in relation to the research and therapy.

1. Introduction

Hearing voices (‘Auditory Verbal Hallucinations’) is commonly reported by people diagnosed with psychiatric conditions such as schizophrenia (Slade and Bentall, 1988), mood disorders (Toh et al., 2015), borderline personality disorder (Slotema et al., 2012), and post-traumatic stress disorder (Butler et al., 1996), as well as in the general (non-clinical) population (Honig et al., 1998). The featural properties of voice hearing (e.g. form, content, impact) appears similar across population groups (Daalman et al., 2011; Slotema et al., 2012), although the emotional and behavioral responses to such experiences vary considerably (Chadwick and Birchwood, 1994; Honig et al., 1998).

Voice content (especially negative content) might not be the only source of individual differences in emotional and behavioral responses to voices (van der Gaag et al., 2003). In their influential cognitive theory, Chadwick and Birchwood (1994) suggested that beliefs held...
about the purpose and intent of voices play a critical role in determining how people respond and cope with these experiences. In particular, they suggested that beliefs about voice omnipotence (i.e. the perceived power of voices) and voice intent (i.e. perceived malevolent or benevolent intentions of voices) can explain the way that voices are responded to, acted on, and complied with (Birchwood and Chadwick, 1997), Chadwick and Birchwood (1994). Specifically, they specified two broad forms of emotional and behavioral responding: Engagement and Resistance. Whilst some people willingly and actively engage with voices, others attempt to resist voices through arguing, shouting or non-compliance (Chadwick and Birchwood, 1994). Evidence presented in support shows that engagement and resistance appear to have differential associations with distress. Engagement is associated with reduced levels of depression and anxiety whilst resistance is positively associated with distress. Engagement and resistance appear to have different numbers of items.

The 30-item Beliefs about Voices Questionnaire (BAVQ) (Chadwick and Birchwood, 1995) and the subsequent 35-item revised version (BAVQ-R) (Chadwick et al., 2000) were designed to measure these key aspects of the voice hearing experience. The BAVQ-R has three belief subscales: one measuring beliefs about Omnipotence, and two intended to measure beliefs about voice intent – Malevolent beliefs (beliefs about negative intent) and Benevolent beliefs (beliefs about positive intent). There are two further subscales designed to measure Engagement and Resistance response styles, each further subdivided into emotional and behavioral modes of expression.

A factor analysis for the original 30-item BAVQ scale was conducted with 60 participants (Chadwick and Birchwood, 1995), but the factor structure of the 35-item BAVQ-R has not yet been examined despite its wide usage in research and clinical settings (Peters et al., 2012; Shawyer et al., 2012; Thomas et al., 2015). If we are to further understand individual differences in emotional and behavioral responding to voices, and the mediating role of beliefs, empirical evidence is needed to confirm the separation and distinctiveness of the constructs being measured. Furthermore, a better understanding of the factor structure of the BAVQ-R is crucial in order to advance scientific inquiry in the field and design more effective therapies. Clearly delineated constructs can be used to refine psychological interventions to specifically target these constructs and to seek out mechanisms that maintain distress.

The aim of the paper was to establish, for the first time, the factor structure of the BAVQ-R using confirmatory factor analysis and principal component analysis in a large sample comprising data from 450 participants from five study centers.

### 2. Method

#### 2.1. Design

The study pooled anonymized and de-identified BAVQ-R data from eight independent studies from five study centers providing an adequately sized sample for factor analysis. Each individual study had received ethical approval from their respective ethics committee. Pooling of data from the different studies into the current sample was done by the second author (KH at the University of Bergen, Norway), and was subsequently approved by the Regional Ethical Committee of Western Norway (REK-Vest #2016/576).

#### 2.2. Participants

Table 1 gives detailed information about the participants in this study. There was a combined total of 450 participants contributing data for the belief subscales, with a mean age of 36.88 years (sd = 11.74) and 46% were female. There were 269 participants contributing data for the response subscales. Participants were resident in the UK, Norway, Australia, the USA, Canada or New Zealand. Diagnosis was not an eligibility criterion given the transdiagnostic nature of voice-hearing and the similarity in form, content and characteristics across diagnostic groups (Daalman et al., 2011).

### 2.3. Measures

#### 2.3.1. Beliefs about Voices Questionnaire – Revised (BAVQ-R) (Chadwick et al., 2000)

The BAVQ-R is a 35 item measure with items rated on a four-point scale from ‘disagree’ (= 0) to ‘strongly agree’ (= 3). Items 1–18 are intended to measure Omnipotence (6 items e.g. “my voice is very powerful”), Malevolence (6 items e.g. “my voice is evil”) and Benevolence (6 items e.g. “my voice wants to protect me”). Items 19–35 are intended to measure responses to voices: Emotional Resistance (4 items e.g. “my voice frightens me”); Behavioral Resistance (5 items e.g. “I tell it to leave me alone”); Emotional Engagement (4 items e.g. “my voice reassures me”) and Behavioral Engagement (4 items e.g. “I listen to it because I want to”). In the original psychometric paper, subscale totals are presented. However, we have chosen to present subscale mean scores (0–3) to allow greater ease of comparison between subscales given that different subscales have different numbers of items.

#### 2.4. Data analyses

Data were available from all 450 participants for the three belief subscales (Omnipotence, Malevolence and Benevolence), and from 269 participants for the two response subscales (Resistance and Engagement). We tested the data set for suitability for factor analyses by doing a Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy test. The KMO returned a value of 0.89, showing our sample is adequate for factor analyses. Separate analyses were conducted with the belief items (analysis 1) and the response items (analysis 2) using normalized varimax rotation of the extracted factors and applying Monte Carlo PCA (parallel criteria analysis) to determine the critical eigenvalue for significant factors (Watkins, 2000). We used AMOS 23.0.0 (build 1607) to conduct confirmatory factor analyses (CFA) to test the models derived from the theory and achieved by the principal component analyses.

For CFA, rules of thumb suggest that the comparative fit index (CFI) should be greater than or close to 0.95 to indicate acceptable fit (Hu and Bentler, 1999). A root mean square error of approximation (RMSEA) value of 0.05 or less is considered a good fit,.08 indicates acceptable fit, and 0.10 or more a poor fit (Browne and Cudeck, 1993). Good fit is also indicated by relative chi-square values ($\chi^2$/df) of less than or equal to 2 and acceptable fit is indicated by values between 2 and 3 (Schermelleh-Engel et al., 2003), although this method lacks robustness to e.g. large sample sizes, non-normality and large correlations between variables (Kline, 2011). The CFA analyses were performed with Maximum Likelihood estimation. We performed Bollen-Stine bootstrapping to control for non-normality in the data and used differences in $\chi^2$ together with the associated difference in degrees of freedom to compare the different models. This is a recommended approach as the compared models are nested.

Finally, derived factors were explored to ascertain differences between women and men (t-tests), relationship with age (Pearson’s correlation), differences between study centers (one-way ANOVA with Bonferroni corrected pairwise comparisons) and differences between studies recruiting only participants with a confirmed diagnosis of a schizophrenia spectrum condition and those studies recruiting mixed diagnostic groups (one-way ANOVA with Bonferroni corrected pairwise comparisons).
Table 1
Participant information by sub-sample.

<table>
<thead>
<tr>
<th>Sub-sample</th>
<th>N</th>
<th>Age/years (sd)</th>
<th>Gender (% female)</th>
<th>Participant country of residence</th>
<th>Education</th>
<th>Participant details and study inclusion criteria</th>
<th>Recruitment strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK 1</td>
<td>180</td>
<td>36.65 (11.06)</td>
<td>62%</td>
<td>UK (57%)</td>
<td>High school (43%)</td>
<td>Inclusion criteria were that participants were aged 18–65 years, were fluent English speakers, had been hearing voices for at least 6 months and had heard voices in the past week. Diagnosis was not an inclusion criterion. Diagnosis given by a psychiatrist was reported by participants as follows: Schizophrenia spectrum (57%), Bipolar Disorder (9%), Borderline Personality Disorder (7%), Other mental health diagnosis (17%), No mental health diagnosis (10%)</td>
<td>Participants took part in an anonymous online study which involved completing a battery of questionnaires.</td>
</tr>
<tr>
<td>Norway 1</td>
<td>32</td>
<td>22.68 (8.25)</td>
<td>31%</td>
<td>Norway</td>
<td>High school (32%) University (14%)</td>
<td>All participants met diagnostic criteria for schizophrenia, schizoaffective disorder, acute and transient psychotic disorder, delusional disorder, drug-induced psychosis or major depressive disorder with psychotic features. Diagnoses were made on the basis of SCID 1 interviews, with the exception of one patient whose diagnosis was based on clinical assessments only.</td>
<td>Participants were recruited from University mental health care institutions, both from inpatient and outpatient treatment units.</td>
</tr>
<tr>
<td>Norway 2</td>
<td>22</td>
<td>28.32 (9.41)</td>
<td>50%</td>
<td>Norway</td>
<td>Data not available</td>
<td>All participants had current auditory verbal hallucinations and an ICD-10 diagnosis of schizophrenia confirmed by a psychiatrist</td>
<td>Participants were recruited through psychiatrists for an MR imaging study.</td>
</tr>
<tr>
<td>Norway 3</td>
<td>20</td>
<td>33.18 (11.31)</td>
<td>33%</td>
<td>Norway</td>
<td>Data not available</td>
<td>All participants had persisting auditory verbal hallucinations and an ICD-10 diagnosis of schizophrenia confirmed by a psychiatrist</td>
<td>Participants were recruited through psychiatric institutions or from participants' psychiatrist local district.</td>
</tr>
<tr>
<td>Norway 4</td>
<td>31</td>
<td>36.58 (10.80)</td>
<td>45%</td>
<td>Norway</td>
<td>High school (52%) University (16%)</td>
<td>Participants had current auditory verbal hallucinations and met ICD-10 diagnostic criteria for schizophrenia (N = 27), schizoaffective disorder (N = 1), or persistent delusional disorder (N = 3)</td>
<td>Participants were recruited through consultant psychiatrists, clinical psychologists and psychiatric nurses from outpatient and inpatient mental health clinics.</td>
</tr>
<tr>
<td>Australia 1</td>
<td>33</td>
<td>34.1</td>
<td>36%</td>
<td>Australia</td>
<td>Mean of 10.82 years (2.06) education</td>
<td>Participants were 18–65 years, had heard auditory verbal hallucinations in the past month, met ICD 10 criteria for a diagnosis of schizophrenia spectrum disorder (confirmed by two psychiatrists from clinical notes review and SCAN interviews)</td>
<td>Participants recruited from inpatient and outpatient community clinics.</td>
</tr>
<tr>
<td>Australia 2</td>
<td>59</td>
<td>39.53 (11.23)</td>
<td>46%</td>
<td>Australia</td>
<td>Data not available</td>
<td>All participants had persisting auditory verbal hallucinations and clinical ICD-10 diagnosis of a schizophrenia-related disorder</td>
<td>Participants recruited via community mental health services to take part in psychological intervention studies.</td>
</tr>
<tr>
<td>UK 2</td>
<td>73</td>
<td>42.57 (11.88)</td>
<td>48%</td>
<td>UK</td>
<td>Data not available</td>
<td>Participants were people experiencing auditory verbal hallucinations and currently accessing mental health services (N = 55) and non-clinical voice hearers, not currently accessing mental health services for these experiences (N = 18)</td>
<td>Clinical participants (N = 55) recruited through NHS mental health services; non-clinical participants (N = 18) were recruited through newspaper advertisements and through a voice hearing conference.</td>
</tr>
<tr>
<td>Total</td>
<td>450</td>
<td></td>
<td>46%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Data from Cole et al. (2017).
* Unpublished data from the Bergen Psychosis 2 project (BP2), Haukeland University Hospital, Bergen, Norway.
* Unpublished data from Fallenberg et al. (2014).
* Unpublished data from the Bergen ERC VOICE project, University of Bergen, Norway.
* Data from Kråkvik, et al. (2013).
* Data from Waters et al. (2003).
* Data from Thomas et al. (2011) and unpublished data from the Voice Exchange project, Melbourne, Australia.
* Data from Sorrell et al. (2010) and unpublished data from Dannahy et al. (2011).
* BAVQ-R items 1–18 (omnipotence, malevolence and benevolence subscales) included, items 19–35 (resistance and engagement subscales) not recorded.
Extraction of principal components, using normalized varimax rotation explores if there was a modest, we conducted a principal component analysis on items 1–18 of the BAVQ-R item means and standard deviations. (Hair et al., 1995).

<table>
<thead>
<tr>
<th>Item</th>
<th>N</th>
<th>Mean (0–3)</th>
<th>sd</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. My voice is punishing me for something I have done</td>
<td>450</td>
<td>1.21</td>
<td>1.21</td>
</tr>
<tr>
<td>2. My voice wants to help me</td>
<td>450</td>
<td>1.14</td>
<td>1.20</td>
</tr>
<tr>
<td>3. My voice is very powerful</td>
<td>450</td>
<td>1.78</td>
<td>1.17</td>
</tr>
<tr>
<td>4. My voice is persecuting me for no good reason</td>
<td>450</td>
<td>1.42</td>
<td>1.21</td>
</tr>
<tr>
<td>5. My voice wants to protect me</td>
<td>450</td>
<td>1.00</td>
<td>1.15</td>
</tr>
<tr>
<td>6. My voice seems to know everything about me</td>
<td>450</td>
<td>2.09</td>
<td>1.09</td>
</tr>
<tr>
<td>7. My voice is evil</td>
<td>450</td>
<td>1.48</td>
<td>1.24</td>
</tr>
<tr>
<td>8. My voice is helping to keep me sane</td>
<td>450</td>
<td>0.76</td>
<td>1.05</td>
</tr>
<tr>
<td>9. My voice makes me do things I really don’t want to do</td>
<td>450</td>
<td>1.04</td>
<td>1.15</td>
</tr>
<tr>
<td>10. My voice wants to harm me</td>
<td>450</td>
<td>1.38</td>
<td>1.26</td>
</tr>
<tr>
<td>11. My voice is helping me to develop my special powers or abilities</td>
<td>449</td>
<td>0.80</td>
<td>1.07</td>
</tr>
<tr>
<td>12. I cannot control my voices</td>
<td>450</td>
<td>1.85</td>
<td>1.17</td>
</tr>
<tr>
<td>13. My voice wants me to do bad things</td>
<td>450</td>
<td>1.25</td>
<td>1.22</td>
</tr>
<tr>
<td>14. My voice is helping me to achieve my goal in life</td>
<td>450</td>
<td>0.73</td>
<td>1.06</td>
</tr>
<tr>
<td>15. My voice will harm or kill me if I disobey or resist it</td>
<td>450</td>
<td>0.87</td>
<td>1.15</td>
</tr>
<tr>
<td>16. My voice is trying to corrupt or destroy me</td>
<td>450</td>
<td>1.36</td>
<td>1.23</td>
</tr>
<tr>
<td>17. I am grateful for my voice</td>
<td>450</td>
<td>0.83</td>
<td>1.11</td>
</tr>
<tr>
<td>18. My voice rules my life</td>
<td>450</td>
<td>1.16</td>
<td>1.14</td>
</tr>
<tr>
<td>19. My voice reassures me</td>
<td>450</td>
<td>0.92</td>
<td>1.13</td>
</tr>
<tr>
<td>20. My voice frightens me</td>
<td>450</td>
<td>1.70</td>
<td>1.21</td>
</tr>
<tr>
<td>21. My voice makes me happy</td>
<td>450</td>
<td>0.72</td>
<td>1.05</td>
</tr>
<tr>
<td>22. My voice makes me feel down</td>
<td>450</td>
<td>1.78</td>
<td>1.16</td>
</tr>
<tr>
<td>23. My voice makes me feel angry</td>
<td>450</td>
<td>1.48</td>
<td>1.19</td>
</tr>
<tr>
<td>24. My voice makes me feel calm</td>
<td>450</td>
<td>0.69</td>
<td>0.99</td>
</tr>
<tr>
<td>25. My voice makes me feel anxious</td>
<td>450</td>
<td>1.90</td>
<td>1.11</td>
</tr>
<tr>
<td>26. My voice makes me feel confident</td>
<td>450</td>
<td>0.73</td>
<td>1.03</td>
</tr>
<tr>
<td>27. I try and stop it</td>
<td>269</td>
<td>1.61</td>
<td>1.25</td>
</tr>
<tr>
<td>28. I try and take my mind off it</td>
<td>269</td>
<td>1.99</td>
<td>1.11</td>
</tr>
<tr>
<td>29. I try and stop it</td>
<td>269</td>
<td>1.73</td>
<td>1.25</td>
</tr>
<tr>
<td>30. I do things to prevent it talking</td>
<td>269</td>
<td>1.63</td>
<td>1.23</td>
</tr>
<tr>
<td>31. I am reluctant to obey it</td>
<td>269</td>
<td>2.01</td>
<td>1.18</td>
</tr>
<tr>
<td>32. I listen to it because I want to</td>
<td>269</td>
<td>0.99</td>
<td>1.19</td>
</tr>
<tr>
<td>33. I willingly follow what my voice tells me to do</td>
<td>269</td>
<td>0.77</td>
<td>1.05</td>
</tr>
<tr>
<td>34. I have done things to start to get in contact with my voice</td>
<td>269</td>
<td>0.77</td>
<td>1.07</td>
</tr>
<tr>
<td>35. I seek the advice of my voice</td>
<td>269</td>
<td>0.84</td>
<td>1.14</td>
</tr>
</tbody>
</table>

3. Results
Table 2 shows the 35 BAVQ-R items along with the item means and standard deviations.

3.1. Belief items factor structure
An initial CFA, based on the original three-factor theoretical model for the belief items (Malevolence, Omnipotence and Benevolence) (Chadwick et al., 2000) was conducted, assuming independent factors. The analysis yielded a poor model fit ($\chi^2/df = 7.24$, $CFI = 0.77$, RMSEA = 0.12) (Gaskin and Lim, 2016). A second analysis, allowing for correlated factors yielded a somewhat better fit between the model and the data ($\chi^2/df = 4.38$, $CFI = 0.88$, RMSEA = 0.087). However, there was strong positive correlation (0.83) between Malevolence and Omnipotence items, and several of the Omnipotence items loaded onto more than one factor. This suggests a two-factor structure as more appropriate, with Malevolence and Omnipotence items loading on a single factor which we label ‘Persecutory Beliefs’.

Since the fit of the original theoretical three-factor model was rather modest, we conducted a principal component analysis on items 1–18 of the BAVQ-R (the items designed to measure beliefs about voices) to explore if there was a different factor structure in the existing data. Extraction of principal components, using normalized varimax rotation of variables, revealed 3 factors with eigenvalue > 1 (eigenvalues: 5.959, 3.443, and 1.224). The Monte Carlo PCA analysis however showed a critical eigenvalue of 1.240 for the third factor, which identified a two-factor solution, explaining 52.23% of the common variance. All of the items from the Malevolence and Omnipotence subscales loaded strongly onto Factor 1. Factor 2 consisted items from the Benevolence sub-scale only. Thus, the current analysis favored a two-factor solution, with a Malevolence/Omnipotence factor (Persecutory beliefs about voices) and a Benevolence factor. There were no items without a unique solution. Internal consistency of the Persecutory and Benevolence subscales were excellent (Cronbach $\alpha = 0.88$ and 0.87 respectively).

A confirmatory factor analysis executed on this new two-factor model however yielded poor goodness of fit ($\chi^2/df = 5.01$, $CFI = 0.86$, RMSEA = 0.094). The still rather poor fit with the data could be explained by cross-loadings in the factor solutions, and also to co-variances across several of the Omnipotence items. We therefore ran a third CFA analysis, and as suggested by the change indexes, we omitted three of the Omnipotence items with the highest co-variances (‘my voice is very powerful’, ‘my voices seems to know everything about me’ and ‘I cannot control my voices’) and one Benevolence item (‘my voice is helping me to develop my special powers or abilities’). This resulted in a better fit for a two-factor model ($\chi^2/df = 4.19$, $CFI = 0.93$, RMSEA = 0.084), see Fig. 1. The difference in $\chi^2$ between the theoretically derived three-factor model and the empirically derived two-factor model was significant ($\chi^2$ diff. (56) = 217.0, $p < 0.001$), showing that the two-factor model was a significantly better fit to the data than the three-factor model. In addition, the final, modified two-factor model show even better fit compared to the theory based three-factor model when independent factors are assumed ($\chi^2$ diff. (61) = 667.80, $p < 0.001$). The Bollen-Stine bootstrap suggested fit with the data both for the original three-factor model as well as for the revised two-factor model (Bollen-Stine bootstrap $p < 0.001$). However, as the average $\chi^2$ for the original three-factor model was 155.2, and the final two-factor model yielded a $\chi^2$ of 91.4, showing that the revised model yielded better fit. In addition, the correlation between the factors was acceptable in this two-factor model ($r = -0.41$) showing each of the factors are explaining a reasonable amount of unique variance. The factor loadings for the final two-factor solution are shown in Table 3, with loadings > 0.40 shown in bold (Hair et al., 1995).

3.2. Response items factor structure
The response items yielded a KMO of 0.895, and as such the sample should be adequate for factor analyses. A confirmatory factor analysis of the response items, assuming independent factors, yielded a poor model fit ($\chi^2/df = 5.20$, $CFI = 0.75$, RMSEA = 0.13). When the analysis was repeated accepting correlated factors, the model yielded a better fit ($\chi^2/df = 2.05$, $CFI = 0.94$ and RMSEA = 0.062). However, the correlation between the two Engagement factors were high (0.91), and also the correlation was also high between the two Resistance factors (0.73). This indicates that the two factors are basically measuring the same concept, and the factors cannot be reliably differentiated at a conceptual level. The change indexes suggested there were cross-variance at several levels. Item 32 (“listen willingly”) in particular showed strong cross loads.

To investigate if another factor structure could be found in the data, we conducted a principal components analysis, using normalized Varimax rotation (Costello and Osborne, 2005) on items 19–35 of the BAVQ-R. The analysis revealed 3 factors with eigenvalue > 1 (eigenvalues: 6.240, 2.678, and 1.168). A Monte Carlo PCA analysis (Ledesma and Valero-Mora, 2007) however suggested a critical eigenvalue of 1.289 for the third factor, which identified a two-factor solution, explaining 52.46% of the common variance. Factor 1 consisted of all of the Resistance items and Factor 2 consisted of all of the Engagement items, with no separation found between Emotional and Behavioral Resistance items on either factor. Thus, the PCA favored a two-factor solution, corresponding to Resistance and Engagement subscales with items measuring emotional and behavioral responding loading together. There were no items without a unique solution. Internal consistency of the Resistance sub-scale was excellent (Cronbach $\alpha = 0.88$ and 0.87 respectively).
consistency of the Engagement and Resistance subscales was excellent (Cronbach $\alpha = 0.87$ and 0.88 respectively). When testing this two-factor solution (Resistance and Engagement) with a CFA, and with control for covariance within the factors, the results yielded a reasonable model fit. However, two behavioral engagement items (‘I listen to it because I want to’ and ‘I willingly follow what my voice tells me to do’) showed strong cross loadings with the other factor. When they were removed from the model, and correlations between variables were controlled for, the analysis suggested a good model fit (the $\chi^2$/df = 2.20, CFI increased to 0.94, and RMSEA = 0.066). Stine-Bollen bootstrapping also confirmed we had reached a reliable model. The modified two-factor model yielded a better fit with the data compared to the theoretically derived four-factor model ($\chi^2$diff. $(33) = 429.3$, $p < 0.001$), and also different from the four factor model with assumed correlated factors ($\chi^2$diff. $(27) = 43.5$, $p < 0.05$), see Fig. 2. In addition, the correlation between the two factors was acceptable ($r = −0.42$), showing each factor has unique contribution to the explained variance. The factor loadings for the final two-factor solution are shown in Table 4, where loadings > 0.40 are shown in bold.

3.3. Demographic and diagnostic associations with newly derived factors

Table 5 shows the newly derived factor means and standard deviations (29 items) by gender, age and sample. Women had significantly higher scores than men on the Persecutory and Resistance factors suggesting that women both hold stronger persecutory beliefs about voices than men and that they are more likely to employ resistance responses. There were no gender differences in Benevolence or Engagement and age was not associated with scores on any of the factors. Finally, Bonferroni corrected pairwise comparisons using one-way ANOVA between the eight samples showed no significant differences in means on the newly derived factors between sites. There were no differences between those studies recruiting only participants with a confirmed diagnosis of a schizophrenia spectrum condition and those studies recruiting mixed diagnostic groups (see Table 1). There were no differences between these two sets of studies on subscale scores of Persecutory beliefs, Benevolence beliefs, Engagement or Resistance.

4. Discussion

The Beliefs about Voices Questionnaire – Revised (BAVQ-R) (Chadwick et al., 2000) was developed to measure key aspects of the voice hearing experience as identified in Chadwick and Birchwood’s
cognitive theory. Whilst their theory suggested that three types of beliefs exist (Omnipotence, Malevolence and Benevolence), the current factor analysis failed to support separation between Omnipo-
tence and Malevolence, with items loading onto a single factor which we have termed Persecutory beliefs and with this requiring removal of three Omnipotence items. Support however was found for Benevolence being a distinct factor.

When it comes to items measuring responding to voices, the use of Resistance and Engagement subscales in the original BAVQ-R was supported in the current analysis following removal of two Engagement items, although we found no support for the separation of emotional and behavioral modes of responding in either subscale. There are a number of theoretical, clinical and research implications of these findings which are explored below.

4.1. Implications

Items measuring omnipotent and malevolent beliefs about voices did not separate but instead correlated strongly with each other. This suggests that voice characteristics of perceived power and perceived hostile intent tend to cluster together and measure the same construct. This finding has also been observed when participants rate their voices on measures of interpersonal relating (Thomas et al., 2009), although it contrasts with the broader literature on social interactions in which power (or dominance) and hostility form two distinct, orthogonal, di-
mensions (Kiesler, 1996). This suggests di-
fferences in the interpersonal relationship of individuals with voices relative to their relationships to people. Perhaps dominance and hostility typically become merged in the context of hallucinated distressing voices. This explanation is also consistent with a developmental perspective which suggests that people distressed by hearing voices are disproportionately likely to have ex-
perienced childhood sexual or physical abuse (Read et al., 2005), and to have been a victim of bullying (Kråkvik et al., 2015). Beliefs about voices are related to broader interpersonal schema (Birchwood et al., 2004; Thomas et al., 2015), which are thought to be shaped by such early life experiences (Beck, 2005). Hence, persecutory beliefs about voices may reflect early experiences of being overpowered by an abuser or bully who intends harm. Indeed, there is growing interest in applying post-traumatic stress disorder theories and interventions to people experiencing psychosis (van den Berg and van der Gaag, 2012) as the relationship between early trauma and later psychotic experiences is becoming more firmly established.

The tendency to respond to voices either by resisting them or by engaging with them emerged as separate factors. However, the se-
paration of emotional and behavioral modes of expression was not supported, corroborating Chadwick’s (2006) suggestion that the
C. Strauss et al.

Table 5
Means and standard deviations on newly derived factors (29 items) by gender, age, site and diagnosis.

<table>
<thead>
<tr>
<th>Gender/mean (sd);</th>
<th>Persecutory beliefs mean (sd) N = 450</th>
<th>Benevolence beliefs mean (sd) N = 450</th>
<th>Resistance mean (sd) N = 269</th>
<th>Engagement mean (sd) N = 269</th>
</tr>
</thead>
<tbody>
<tr>
<td>females</td>
<td>1.23 (0.93)</td>
<td>0.80 (0.86)</td>
<td>2.05 (0.81)</td>
<td>0.70 (0.87)</td>
</tr>
<tr>
<td>males</td>
<td>1.10 (0.81)</td>
<td>0.89 (0.90)</td>
<td>1.58 (0.86)</td>
<td>0.86 (0.86)</td>
</tr>
<tr>
<td>t (p)</td>
<td>2.63 (0.009)</td>
<td>1.06 (0.29)</td>
<td>4.08</td>
<td>1.44 (0.15)</td>
</tr>
<tr>
<td>correlation with age/r (p)</td>
<td>-0.09 (0.09)</td>
<td>0.03 (0.59)</td>
<td>-0.14 (0.049)</td>
<td>0.1 (0.15)</td>
</tr>
</tbody>
</table>

Sub-samples by site/mean (sd):

- UK 1a: 1.10 (0.86) 0.81 (0.88) 1.78 (0.73) 0.78 (0.77)
- Norway 1b: 1.30 (0.81) 0.92 (0.86) 1.64 (0.78) 0.85 (0.72)
- Norway 2c: 1.17 (0.81) 0.57 (0.59) 1.70 (0.73) 0.58 (0.54)
- Norway 3d: 1.72 (0.71) 1.08 (0.91) 2.04 (0.54) 0.87 (0.87)
- Norway 4e: 1.45 (0.78) 0.77 (0.87) 1.81 (0.65) 0.58 (0.78)
- Australia 1f: 1.19 (0.59) 1.15 (0.51) 1.54 (0.83) 0.81 (0.65)
- Australia 2g: 1.20 (0.79) 1.00 (0.90) 1.85 (0.75) 0.93 (0.84)
- UK 2h: 1.42 (1.06) 0.88 (0.99) 1.83 (1.10) 0.80 (1.01)

Bonferroni-corrected pairwise comparisons

Sub-samples by diagnosis/mean (sd):

- Schizophrenia-spectrum only studies: 1.30 (0.76) 0.94 (0.81) 1.78 (0.73) 0.78 (0.77)
- Mixed-diagnoses studies: 1.21 (0.92) 0.84 (0.90) 1.77 (1.01) 0.82 (0.93)
- t (p) = 2.22 (0.22) 1.16 (0.25) 0.09 (0.93) -0.31 (0.76)

Notes:

1. Unpublished data from Relationships and Recovery in Psychosis project (RRP), University of Surrey, UK.
2. Unpublished data from the Bergen Psychosis 2 project (BP2), Haukeland University Hospital, Bergen, Norway.
4. Unpublished data from the Bergen ERC VOICE project, University of Bergen, Norway.
5. Data from Kråkvik et al. (2013).
6. Data from Waters et al. (2003).
7. Data from Thomas et al. (2011) and unpublished data from the Voice Exchange project, Melbourne, Australia.
8. Data from Sorrell et al. (2010) and unpublished data from Dannahy et al. (2011).
9. Significant Bonferroni corrected t-tests (between women and men; between studies recruiting only people with schizophrenia-spectrum diagnosis and studies recruiting mixed-diagnosis samples); Bonferroni corrected Pearson’s r correlations (with age); and one-way ANOVA Bonferroni corrected pairwise differences (between study sites).
10. It is of note that resistant responding to voices is associated with anxiety and depression (Chadwick et al., 2000) and this has important clinical implications. ‘Third-wave’ cognitive behavior therapies emphasize the potentially counterproductive effects of attempting to resist unpleasant experiences. Rather than trying to resist voices, mindfulness-based approaches and Acceptance and Commitment Therapy encourage non-judgmental acceptance of voice hearing experiences (Chadwick et al., 2016; Dannahy et al., 2011; Shawyer et al., 2012; Strauss et al., 2015; Thomas et al., 2014) and so lend themselves well as therapeutic approaches to reducing voice resistance.

The original BAVQ-R study was restricted to people with diagnoses on the schizophrenia spectrum (Chadwick et al., 2000). Diagnosis was not an eligibility criterion for this study. There were no differences in ratings for Persecutory beliefs, Benevolence beliefs, Engagement or Resistance between those studies recruiting only participants with a confirmed diagnosis of a schizophrenia spectrum condition and studies recruiting mixed diagnostic groups. This finding adds weight to the hypothesis of the transdiagnostic nature of voice hearing and the similarity in form, content and characteristics across diagnostic groups (Daalman et al., 2011).

It is important to note that these findings apply to group-level data and that individual experiences will vary. Some people, for example, will experience voices as both omnipotent and benevolent. Whilst at the group-level omnipotence items and malevolence items loaded together, this does not mean that everyone who experiences voices as omnipotence will also experience them as malevolent. This underlines the need for careful assessment and formulation when working therapeutically and not to make assumptions based on group-level findings.

4.2. Strengths and limitations

A strength of this study is that it included a sufficiently large sample to conduct principal components analysis. Indeed, with 450 participants for the beliefs items, this is the largest BAVQ-R dataset that has been reported in the literature to our knowledge. The large sample also allowed for sub-sample differences to be explored regarding gender and age that could be taken forward and explored in more detail in future studies.

There are potential limitations with BAVQ-R items. Two omnipotence items imply malevolent voice intent (“my voice makes me do things I really don’t want to do” and “my voices will harm or kill me if I disobey or resist it”) which may explain in part why these items loaded together. Future research could develop omnipotence items which do not imply either malevolent or benevolent intent in order to further explore the underlying factor structure of a measure. Also, the samples did not differ from each other on any of the newly derived subscales, so, although the samples were different demographically they were not different in terms of beliefs about voices or in ways of responding to voices. Moreover, there were no differences in subscale scores between the studies where all participants had a confirmed diagnosis of a schizophrenia spectrum condition and those studies recruiting mixed diagnostic groups.

There are potential limitations with BAVQ-R items. Two omnipotence items imply malevolent voice intent (“my voice makes me do things I really don’t want to do” and “my voices will harm or kill me if I disobey or resist it”) which may explain in part why these items loaded together. Future research could develop omnipotence items which do not imply either malevolent or benevolent intent in order to further explore the overlap found in the current study between omnipotence and malevolence. The final models provided the best fit to the data amongst the models tested, with fit statistics being in the acceptable range for the final derived two belief factors and two response factors. To further aid the measurement of the proposed constructs in a more precise and coherent way, future research could aim to improve model fit by developing new items. However, the BAVQ-R is widely used and no other more psychometrically robust measure of beliefs of voices exists to our knowledge. The current study provides empirical support for using a 29-item version of the BAVQ-R in clinical practice and in research settings. There were a few missing data from some studies for
some of the BAVQ-R items and in some cases full demographic information was not available. The eight studies also did not use a consistent set of additional measures to allow convergent and divergent validity of the newly derived factors to be tested in the full sample of 450 participants. However, because the overall sample was so large the analyses with missing data were still adequately powered to detect medium effect sizes.

There is growing interest in third-wave CBT approaches for psychosis which focus on the relationship with psychotic experiences such as hearing voices, and emphasize the role of mindfulness, acceptance, self-compassion and values-directed action. It could be argued that as hearing voices, and emphasize the role of mindfulness, acceptance, coherence which focus on the relationship with psychotic experiences such as hearing voices.

Both effects on the primary outcome (compliance) (Birchwood et al., 2011), suggesting that these approaches are still worthy of implementation and further investigation. We suggest therefore that constructs central to second-wave CBT for distressing voices – namely beliefs about voices – still require valid and reliable methods of assessment.

4.3. Conclusions

A factor structure for the BAVQ-R is suggested where Malevolence and Omnipotence items are combined into a single Persecutory beliefs factor, and whilst Resistance and Engagement subscales were confirmed, further separation of these into behavioral and emotional subscales is not warranted from the data. We suggest that the 29 BAVQ-R items contributing to the final factor structure could be used in future clinical practice and research. This revised factor structure will hopefully facilitate future research exploring the mechanisms involved in causing voice-distress and when evaluating therapeutic change mechanisms.

Acknowledgements

We would like to thank all the participants who took part in the studies that contributed to this paper. In addition, we would also like to thank the researchers who helped to collect the data for the studies included in this paper: Alison Holt, Esther Cole and Eleanor Sorrell (University of Surrey, UK); Jill Bjarke, Lena A. Stabell, Eirik Kjelby and Farivar Fathian (Bergen, Norway); Rolf Wilhelm Gråwe (St. Olav’s University Hospital, Trondheim, Norway); or who contributed to earlier studies in which the data were collected: Johanna Baddock, Murray Mayberry and Patricia Mitchie (University of Western Australia, Australia).

Conflicts of interest

None.

Funding

Studies included in this paper received funding as follows: Norwegian Research Council (#213727) and the Western Norway Regional Health Authority (# 911820) grant to Erik Johnsen; Research Council of Norway (RCN) Grants #190676/V50 and #223273 to Kenneth Hugdahl; and European Research Council (ERC) Advanced Grants #249516 and #693124 to Kenneth Hugdahl.

References


