

Reality of auditory verbal hallucinations

Tuukka T. Raij,^{1,2} Minna Valkonen-Korhonen,³ Matti Holi,⁴ Sebastian Therman,⁵
Johannes Lehtonen³ and Riitta Hari^{1,6}

- 1 Brain Research Unit, Low Temperature Laboratory and Advanced Magnetic Imaging Centre, Helsinki University of Technology, Espoo, Finland
- 2 Department of Psychiatry, University of Helsinki, Helsinki, Finland
- 3 Department of Psychiatry, University of Kuopio and University Hospital of Kuopio, Kuopio, Finland
- 4 Kellokoski Hospital, Kellokoski, Finland
- 5 Department of Mental Health and Alcohol Research, National Public Health Institute, Helsinki, Finland
- 6 Department of Clinical Neurophysiology, Helsinki University Central Hospital, Helsinki, Finland

Correspondence to: Tuukka T. Raij, MD, PhD,
Advanced Magnetic Imaging Centre,
Box 3000,
Helsinki University of Technology,
Finland
E-mail: rajitu@neuro.hut.fi

Distortion of the sense of reality, actualized in delusions and hallucinations, is the key feature of psychosis but the underlying neuronal correlates remain largely unknown. We studied 11 highly functioning subjects with schizophrenia or schizoaffective disorder while they rated the reality of auditory verbal hallucinations (AVH) during functional magnetic resonance imaging (fMRI). The subjective reality of AVH correlated strongly and specifically with the hallucination-related activation strength of the inferior frontal gyri (IFG), including the Broca's language region. Furthermore, how real the hallucination that subjects experienced was depended on the hallucination-related coupling between the IFG, the ventral striatum, the auditory cortex, the right posterior temporal lobe, and the cingulate cortex. Our findings suggest that the subjective reality of AVH is related to motor mechanisms of speech comprehension, with contributions from sensory and salience-detection-related brain regions as well as circuitries related to self-monitoring and the experience of agency.

Keywords: brain; functional magnetic resonance imaging; reality distortion; auditory verbal hallucination; inferior frontal gyrus

Abbreviations: AVH = auditory verbal hallucination; IFG = inferior frontal gyrus; HRF = haemodynamic response function; PANSS = Positive and Negative Syndromes Scale; SRH = subjective reality of hallucinations; VAS = visual analogue scale

Introduction

When John Nash was asked how he, 'a mathematician, a man devoted to reason and logical proof', could believe that extraterrestrials were sending him messages, he answered 'because the ideas I had about supernatural beings came to me the same way that my mathematical ideas did. So I took them seriously' (Nasar, 1998). Besides delusions, auditory verbal hallucinations (AVHs) may appear very real to the subject for reasons that remain largely unknown. The strong subjective reality of AVHs may result in inappropriate behaviour which, in extreme cases, can be life-threatening.

The subjective reality of hallucinations (SRH) is most probably related to the perceptual characteristics of AVHs (Hunter, 2004). However, we refer to the SRH, assessed as an experience of 'voices' on a continuum from imaginary or unreal to real, as a broader concept that includes the salience of AVH (Kapur, 2003) as well as compromised self monitoring (Frith, 1992), meaning impaired ability to recognize one's own mental functioning. Frith and Done (1989) suggested that compromised self-monitoring is associated with the experience that AVHs are beyond one's control, a phenomenon likely intensifying the SRH. Furthermore, the SRH seems to be related to a shift from one's own to an alien

Received April 2, 2009. Revised May 28, 2009. Accepted June 7, 2009. Advance Access publication July 20, 2009

© The Author(s) 2009. Published by Oxford University Press on behalf of Brain.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/2.5/uk/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

agency, so that, for example, self-generated verbal material is experienced as originating from a non-self author (Frith, 2005).

Early studies on the brain correlates of reality distortion compared baseline brain metabolism in subjects with and without reality distortion symptoms (delusions and hallucinations) (Liddle *et al.*, 1992; Kaplan *et al.*, 1993). More recent studies have focused on memory errors when the subjects attempt to remember whether words in a sentence were imagined or perceived, or on emotional picture processing in patients with and without delusions and hallucinations (Surguladze *et al.*, 2006; Taylor *et al.*, 2007; Vinogradov *et al.*, 2008). Imaging studies have also compared brain activity during hallucination versus non-hallucination periods, identifying wide-spread brain activation related to AVHs (Tiihonen *et al.*, 1992; Silbersweig *et al.*, 1995; Dierks *et al.*, 1999; Lennox *et al.*, 2000; Shergill *et al.*, 2000; van de Ven *et al.*, 2005; Hoffman *et al.*, 2007; Sommer *et al.*, 2008). These activations frequently include the inferior frontal gyri (IFG), the anterior cingulate cortex, the parahippocampal gyrus, and the superior and middle temporal gyri; for a review, see Allen *et al.* (2008). Because these studies did not rate the SRH, and because AVHs are coupled with multiple cognitive and emotional functions, it is difficult to determine which of the observed brain activations would be associated with key characteristics of AVHs (Woodruff, 2004). No previous studies have directly assessed the connection between the distortion of reality and the brain function during delusions or hallucinations, probably because finding subjects suitable for and cooperative with such a study is difficult. Optimal subjects should experience recurring symptoms that vary sufficiently in subjective reality during brain scanning. Furthermore, the subjects should be able to rate this dimension reliably, even if cognitive processing is frequently compromised in disorders with reality distortion.

In the present study, we were fortunate to work with 11 highly functioning subjects who experienced intermittent AVHs during functional magnetic resonance imaging (fMRI) and were able to rate the subjective reality of their AVHs reliably (on a continuum from imaginary to real voices). Based on these data, we first identified brain correlates of the SRH in AVH-related brain activation (Fig. 1). Because many psychiatric disorders (Andreasen, 1997), particularly schizophrenia (Friston, 2002), have been related to distorted interaction within large-scale neuronal circuitries, we next quantified coupling between the brain correlates of the SRH and other brain regions with an established method of psychophysiological interaction (Friston *et al.*, 1997). Finally, we correlated this coupling with the subjects' SRH ratings (see 'Statistical methods' section for *a priori* regions).

Materials and methods

Subjects and pre-examinations

With the help of a Finnish voice hearers' association and local mental health personnel, we delivered 200 letters to psychiatric outpatients and third-sector association members known to experience AVH. In the letter, we described the study and asked subjects to contact us if they believed they would hear several intermittent voices, 10–60 s in duration, during a 30-min noisy fMRI scanning. Inclusion further required

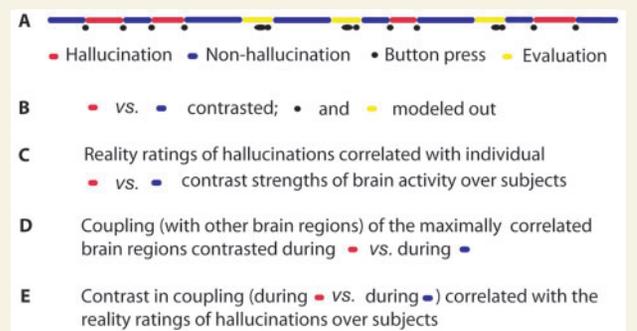


Figure 1 During fMRI scanning, the subjects signalled the beginning and end of multiple AVHs and rated the SRH and loudness of AVHs by pressing two buttons. (A) Schematic example of a part of a session; time runs from left to right. Contrast images for hallucination versus non-hallucination periods (B) were correlated with the SRH over subjects (C). We then compared the coupling of the so-found neuronal correlates of the SRH between hallucination versus non-hallucination periods (D). Finally, we correlated the hallucination-related coupling with the SRH (E).

the lack of neurological or severe somatic disorders, and no contraindications of MR imaging, including severe obesity. Of the 43 subjects who replied, those 31 who were most likely to fulfil our inclusion criteria received a detailed questionnaire about voices and health based on the literature on AVH (Hustig and Hafner, 1990; Oulis *et al.*, 1995; Nayani and David, 1996; David, 1999; Stephane *et al.*, 2003).

To test the reliability of the answers, this questionnaire included, for the SRH, two visual analogue scales (VAS) identical in content but slightly different in wording: For the first ('Are the voices more imaginary or real?'), the endpoints were labelled 'imaginary' and 'real', and for the second ('How real are the voices?'), the endpoints were 'not real at all' and 'real'.

Of the 30 respondents to the questionnaire, we included 13 on the basis of their cooperation and the above-mentioned inclusion criteria. The most frequent contraindication was having body dimensions (reported by the patients in the questionnaire) above the limits of the scanner. One subject required several repetitions of the instructions during rehearsal of the fMRI part and had difficulty learning the task. We therefore excluded him from the scanning. Another subject could not enter the scanner due to claustrophobia.

For the remaining 11 subjects (six males, five females; mean age 35 years, range 23–52 years, 10 right- and 1 left-handed), an experienced psychiatrist (M.H.) conducted a diagnostic interview (American Psychiatric Association, 1994) and Positive and Negative Symptom Scale (PANSS) assessment (Kay *et al.*, 1987). An experienced clinical psychologist conducted cognitive testing, including the Finnish version of the Wechsler Adult Intelligence Scale, 3rd Edition (Wechsler, 2007).

In addition, subjects filled the Beck Cognitive Insight Scale (Beck *et al.*, 2004). This scale has two subscales that address cognitive responses to psychotic symptoms. One subscale measures self-reflectivity and another self-certainty. The self-certainty subscale relies on findings that psychotic patients tend to jump to conclusion without considering alternative explanations: it consists of a 1–4-point scale of agreement with statements such as 'My interpretations of my experiences are definitely right'.

The study received the prior approval of the local ethics committee and each subject who returned the questionnaire about voices and health also provided an informed written consent form.

fMRI

Before entering the fMRI scanner, subjects practiced the task on a computer with type token ratio (T.T.R.) until they completely understood and were able to perform it.

We collected functional whole-brain images by measuring the blood oxygenation-dependent (BOLD) signal [Signa VH/i 3.0T MRI scanner; GE Healthcare, Chalmont St Giles, UK; echo time (TE) 2 ms, repetition time (TR) 2.3 s, flip angle 75°, field of view (FOV) 24 cm, 39 oblique slices aligned with the anterior-posterior commissure line, slice thickness 4 mm and matrix size 64 × 64]. In addition, T₁-weighted structural images were collected for each subject.

During fMRI sessions, subjects had cylinder-shaped response keys in both hands. They indicated each beginning and each end of the intermittent AVH with a short button press, using either the left or right thumb. To compensate for the small number of suitable subjects, we collected fMRI data during a large number (altogether 585) of AVHs. The first four whole-head images of each session were automatically discarded to allow stabilization of the T1 effect.

To avoid any confusion between the beginnings and ends of the AVH, we presented visual feedback (in Finnish) on a projector screen using Presentation® software (Version 0.70, <http://www.neuro-bs.com>): 'Voices present—please push any button when the voices stop', or 'Voices absent—please push any button when the voices begin', respectively. Anytime an 18-s period passed without hallucination, a 100-point VAS appeared asking the subject to rate alternately the reality or loudness of the latest AVH with the questions 'How real were the voices?' (endpoints 'imaginary' and 'real') or 'How loud were the voices?' (endpoints 'just audible' and 'loudest possible'). The subjects moved a cursor to the left or right by depressing the left- or right-hand button, respectively; the answer was registered 3 s after the subject stopped moving the cursor. The 18 s delay before evaluation was considered necessary and sufficient for collecting post-hallucination baseline data for comparison with the hallucination-related activation.

If the hallucination-free period continued for 18 s after the first evaluation, the other VAS appeared, followed by the text: 'Voices absent—please push any button when the voices begin'. This text appeared until the end of the session or until the subject signalled the beginning of a new AVH.

The subjects participated in one to six (mean 4) scanning sessions (256–512 images during each 10–20 min period). After each session, the subject rated the mean SRH and other subjective dimensions (Supplementary Table 2) of the AVH during the past session (post-session ratings). After the entire scanning session, the subjects also evaluated their accuracy in VAS scaling and signalling the beginnings and ends of the AVH; eight subjects reported having correctly signalled all AVHs, and three subjects reported having correctly signalled most AVH. Of the 11 subjects, 10 evaluated their task performance, including ratings, as either good or rather good. One subject experienced only a few non-hallucination periods that were too short for intra-session rating; for this subject, only post-session ratings were included in the analysis.

The site of the auditory cortex was identified in eight subjects by presenting, in a separate 8-min session, speech and brief sounds (0.1 s tones of 250, 500, 1000, 2000 or 4000 Hz, repeated in random order at 5 Hz; eight 30-s blocks of stimulation, each separated by a 30-s rest period). The stimuli were presented at 40 dB above the individual's hearing threshold.

Analysis of subjective evaluations

We correlated the mean post-session ratings of the SRH with the means of other subjective dimensions of AVHs and with Beck cognitive

insight scores across subjects. The statistical significance of the correlations was addressed with the Pearson's test.

Analysis of hallucination-related brain activation

The functional MR images were preprocessed with established methods in statistical parametric mapping (SPM)-2 software (<http://www.fil.ion.ucl.ac.uk/spm/spm2.html>). For the first analysis, we created individual boxcar regressors for hallucination and non-hallucination periods (without evaluation or any other task), button presses (with duration of 2.3 s) and evaluation periods, convolved with a haemodynamic response function (HRF). High-pass filtering was applied according to the temporal variation of AVH periods (cut-off 128–300 s), and a first-order autoregressive model was included to compensate for any autocorrelation error (Bullmore *et al.*, 1996). Movement regressors were included in the model whenever movement exceeded 1 mm in any direction.

First, we contrasted the hallucination periods and the non-hallucination periods. To avoid contamination of the fMRI results with the signalling of the beginnings and endings of the hallucination periods, both hallucination and non-hallucination periods required equal preparation for the button press and button press-related activation was regressed out in the analysis. In addition, evaluation period-related activation was regressed out before the individual contrast images were created for hallucination versus non-hallucination periods. We entered the resulting individual contrast images into the one-sample *t* test to reveal AVH-related activation at the group level. We then correlated the individual contrast images voxel-wise with the average SRH ratings across subjects. Preliminary correlation analysis with intra-individual variance of the SRH resulted in no statistically significant findings, probably due to the low signal-to-noise ratio. Therefore, to minimize the effect of random variation, we conducted correlation analysis over subjects with carefully calculated mean SRH values. These SRH values were achieved by first calculating the mean of all the SRH ratings in a single session, then computing the average of this mean value and the post-session rating (considering all AVHs during the session), and finally calculating the mean of all averages across all the sessions for each subject. Also the post-session ratings were used because only mean 4 of 16 single AVHs were rated per session.

After correlating AVH-related activations voxel-wise with the SRH, we entered—in a series of independent analyses—three potential confounders into the correlation analysis: To rule out the effects of medication-related dopamin-2-receptor blocking on neurovascular coupling, and therefore on the fMRI signal, we added chlorpromazine-equivalent doses of any anti-psychotic medication (Centorrino *et al.*, 2002) as a confounding covariate. Because the cognitive style of interpretation could interfere with subjective reporting of the SRH, the self-certainty ratings that correlated with the SRH were added to the analysis as a confounding covariate.

To further test specificity of the correlation between the SRH and the strength of brain activation during AVH, we added the mean loudness estimates, rated and calculated in the same way as the SRH estimates, to the analysis as a confounding factor.

Analysis of hallucination-specific coupling between the IFG and other brain regions

In another analysis, we compared the coupling of the signal from the IFG with other brain regions, between hallucination and

non-hallucination periods. For this analysis, we used the SPM2 psychophysiological interaction tool (Friston *et al.*, 1997) that compares the context-specific (here, AVH-specific) contribution of brain regions to each other, referred to as 'coupling'. Following the procedures of previous studies (Macaluso *et al.*, 2000; Stephan *et al.*, 2003; Pasley *et al.*, 2004; Bingel *et al.*, 2007), we extracted fMRI signals from the seed regions (one for the left and another for the right IFG) for each session in which AVH were signalled. The seed regions were spheres with an 8-mm radius, including 33 voxels, and with their centres at $-52, 20, 8$ in the left IFG, and at $52, 12, 12$ in the right IFG (x, y, z coordinates in the Talairach system); the coordinates were found by visually estimating the centre of gravity in the area of maximum group-level correlation of AVH-related activation with the SRH.

Because coupling between brain regions occurs at the level of neuronal signalling rather than at the level of haemodynamics (Gitelman *et al.*, 2003), we deconvolved the fMRI time courses to estimate the neuronal signal without the haemodynamic lag. We then created a regressor for hallucination versus non-hallucination periods, and multiplied it by the deconvolved time course from the IFG. We ran a new general linear model (GLM) analysis using this regressor, mean-normalized and reconvolved with the HRF. We added the box-car regressor for hallucination versus non-hallucination periods, convolved with the HRF, to the model to remove the AVH-related activation. Therefore, the results reflect coupling during the activation rather than the AVH-related activation. Button-press regressors and movement regressors were applied as in the activation analysis, and the original fMRI-time course from the IFG was added to remove the coupling beyond the contrast of interest. Therefore, the model specifically tested for the contrast of coupling of the IFG with other brain regions during AVH versus non-AVH periods. One-sample t test was used to test resulting contrast images, at the group level, for hallucination-related changes in the connectivity of the IFG with other parts of the brain. Finally, we correlated these contrast images voxel-wise with the SRH across subjects.

We reanalysed the data to exclude contribution of possible global confounds that could arise from respiration or other movements. In the re-analysis, carried out at the individual level in the same way as was the first psychophysiological interaction analysis, we normalized the global mean values by scaling the average voxel value to zero. These normalized contrast images of AVH-specific coupling were then used for second correlation analysis with the SRH.

Statistical methods

Statistically significant activation refers to anatomically meaningful activation in the *a priori* regions with $P < 0.005$ for each of >20 contiguous voxels (corresponding to $P < 0.0007$, uncorrected), or for the correlation with coupling, $P < 0.005$, uncorrected, at the cluster level (based on single voxel values and the extent of activated voxels according to random field theory). We selected this latter liberal threshold to avoid false negative findings in testing the hypothesis about involvement of the large-scale neuronal network. Note, however that most findings are of much higher statistical significance. For the AVH-related brain activations, the *a priori* regions included those repeatedly activated during the AVH in previous studies (Allen *et al.*, 2008): the IFG, the anterior cingulate cortex, the parahippocampal gyrus, and the superior and middle temporal gyri. For the correlation between the SRH and coupling, additional *a priori* regions included those brain areas associated with the factors believed to contribute to reality distortion: the posterior parietal cortices related to self-monitoring and the experience of agency (Jeannerod and Pacherie, 2004; Frith, 2005; Allen *et al.*, 2007), the right posterior temporal

lobe implicated in agency (Tankersley *et al.*, 2007), the anterior and posterior cingulate cortices related to self-monitoring (Northoff and Bermpohl, 2004) and agency (Tomlin *et al.*, 2006), and the striatum involved in the subjective saliency of percepts and in antipsychotic medication (Kapur, 2003; Agid *et al.*, 2007).

To define correlation coefficients, we split data into two independent sets (Kriegeskorte *et al.*, 2009): we gave a serial number to each of the 10 subjects who participated in two or more fMRI sessions, and to each subject's fMRI session. For the first data set, we created contrast images of sessions with an odd number for the subjects with odd number and of sessions with even number for the subjects with even number. For the second data set the odds and evens were reversed. We calculated the individual mean SRH values for both data sets as described above and conducted two independent correlation analyses both with activation and coupling. We defined voxels of maximum correlation ($P < 0.05$, uncorrected) in the regions of interest from one data set and extracted r -values from these voxels from the other data set.

Any difference between correlations of the SRH with the IFG and the temporal lobe activation was tested using a z test (Tabachnick and Fidell, 2001). Because we compared correlations rather than slopes of regression lines, the results indicate whether inter-individual variation of AVH-related brain activation explains the SRH significantly better in one than in another brain region. The functional images were overlaid on an SPM template and a MATLAB program was used to convert MNI (Montreal Neurologic Institute) coordinates to the Talairach system (MNI2TAL, author M. Brett, <http://www.mrc-cbu.cam.ac.uk/Imaging/Common/mnispace.shtml>).

Results

Subject characteristics

Seven subjects had schizophrenia and the remaining four had a closely related schizoaffective disorder, as defined in the Diagnostic and Statistical Manual of Mental Disorders IV (1994). The subjects matched the normal population in cognitive performance (Supplementary Table 1), lived in the community and four subjects worked full time. Except for hallucination evaluations, the subjects' PANSS (Kay *et al.*, 1987) scores were low (mean total score 56, range 40–85; Supplementary Table 1). The two VASs of the SRH, presented with different wordings, correlated with each other ($r = 0.89$, $P < 0.001$) and written descriptions of the hallucinations were coherent. All subjects reported hearing words or sentences that no one else heard. They attributed these experiences to an unknown origin or to their psychiatric disease. None of the subjects reported supernatural beliefs about voices, although many of them had held such beliefs before. Six subjects believed that the AVH had a meaning of their own: to comment, to contact, to accuse, to punish, or to tell them what to do.

Correlations between subjective reality and other subjective dimensions of hallucinations

The mean SRH was similar in subjects who did and did not believe AVHs to have a meaning of their own (43 and 48 of 100,

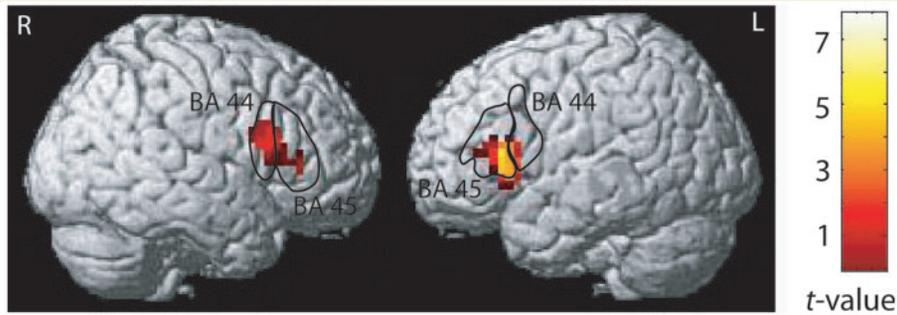


Figure 2 Correlation between AVH-related brain activation and the SRH. Brodmann areas 44 and 45 (the Broca's region and its right homologue) are marked according to cytoarchitectonical maps by Eickhoff *et al.* (2005). Right: the colour scale for statistical significance.

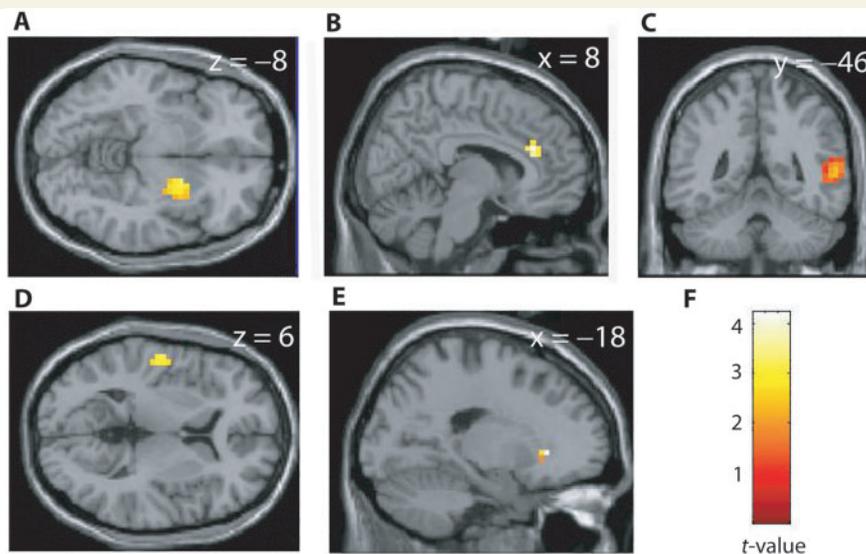


Figure 3 Positive correlations between hallucination-related coupling (in parameter estimates) and the SRH. The coupling refers to the difference with the left IFG ($x, y, z = -52, 20, 8$) during hallucinations versus non-hallucination periods. The stronger the SRH, the stronger the coupling of the IFG with the right ventral striatum (A), the middle right anterior cingulate cortex (B), the right posterior temporal lobe (C), the auditory cortex (D), and the left nucleus accumbens (E). (F) The colour scale for statistical significance.

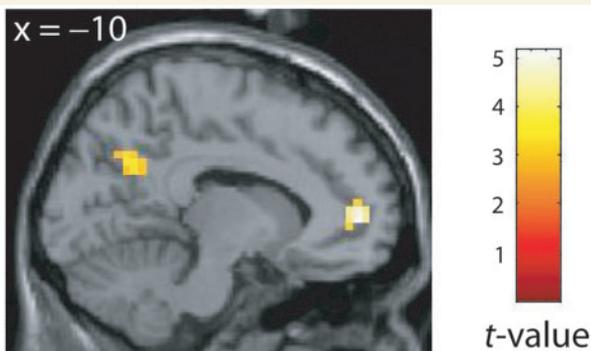


Figure 4 Negative correlation between the AVH-related coupling (in parameter estimates) of the left IFG and the SRH. The stronger the SRH, the weaker the coupling of the IFG with the posterior and rostral anterior cingulate cortex.

respectively). In the post-session ratings (Supplementary Table 2), the SRH correlated positively across subjects with estimates of speech-likeness (versus thought-likeness; $r = 0.54$, $P = 0.04$; Pearson's one-tailed test) and loudness ($r = 0.52$, $P = 0.05$) of the hallucinations, as well as with hallucination-related suffering ($r = 0.59$, $P = 0.03$). The SRH ratings did not correlate with the self-reflectivity subscale of the Beck Cognitive Insight Scale (Beck *et al.*, 2004), but they correlated negatively with the subscale on the self-evaluated certainty of the experiences ($r = -0.74$, $P = 0.02$). Means of single AVH ratings correlated strongly with post-session ratings ($r = 0.90$, $P < 0.001$).

Correlation of AVH-related brain activation with the SRH

On average 53 hallucination periods (range 7–153, mean duration 41 s, range 2–393 s) and 85 non-hallucination periods

(range 10–162, mean duration 23 s, range 1–365 s) occurred per subject during the fMRI recordings. Individual average SRH ranged from 9 to 86 (Supplementary Table 1), and AVH-related brain activation resembled activations observed in previous studies, including the right parahippocampal cortex, the bilateral IFG, the right posterior temporal lobe, the left anterior temporal lobe, and the right anterior cingulate cortex (Supplementary Fig. 1 and Supplementary Table 3).

The strength of AVH-related activation in the IFG, corresponding to the Broca's region and its right-hemisphere homologue, correlated with the SRH across subjects ($P < 0.001$, $r = 0.63$ and $P < 0.001$, $r = 0.73$ for the whole group, and $P = 0.001$ and $P = 0.013$ for the 7 subjects with schizophrenia, respectively; values for the most significantly correlated voxels reported throughout; Fig. 2 and Supplementary Table 4). This correlation was statistically significant ($P < 0.05$) both with and without confounding covariates that included dopamine-2-receptor blocking medication, self-certainty ratings, and the subjective loudness of the AVH. The strength of the AVH-related IFG activation showed a trend towards explaining better ($z = 1.2$, $P = 0.12$) the individual SRH than did the strength of any AVH-related activation in the temporal lobe, including the auditory cortices.

Correlation of the hallucination-specific coupling of the IFG with the SRH

At the group level, no statistically significant differences were found in the coupling of the IFG with other brain regions during AVH-periods versus non-AVH periods ('AVH-related coupling', $P > 0.005$ in the regions of interest, uncorrected, or $P > 0.05$ in other brain regions, corrected for multiple comparisons).

The SRH variability explained the inter-individual variation of the AVH-related coupling: The SRH scores correlated positively with the AVH-related coupling between the left IFG and the following brain regions: the bilateral supratemporal auditory cortex ($P = 0.004$, $r = 0.40$ for the left and $P = 0.024$, $r = 0.34$ for the right hemisphere), the right posterior temporal lobe ($P = 0.008$, $r = 0.24$), the middle right anterior cingulate cortex ($P = 0.001$, $r = 0.18$), the right ventral striatum and the left nucleus accumbens ($P = 0.004$, $r = 0.15$ and $P = 0.001$, $r = 0.21$, respectively; Fig. 3, Supplementary Table 5). The correlation of the SRH with the coupling between the left IFG and the left auditory cortex was strongest in the Heschl's gyrus, within 1 cm of the maximum activation elicited by brief tones in the same subjects ($x, y, z = -51, -15, 8$ versus $-55, -12, 1$, respectively). The SRH values correlated positively also with the coupling of the right IFG with the right posterior superior temporal gyrus ($x, y, z = 59, -26, 26$; $P < 0.001$). These positive correlations were statistically significant both in the entire subject group and in the subgroup with schizophrenia (Supplementary Table 5) and they remained significant with the application of global normalization in the reanalysis ($P < 0.05$).

The SRH values correlated negatively with the AVH-related coupling between the left IFG and the left pregenual and posterior cingulate cortex ($P < 0.001$, $r = -0.57$ and $P = 0.003$, $r = -0.65$, respectively; Fig. 4, Supplementary Table 6) as well as with the AVH-related coupling between the right IFG and the left

pregenual cingulate cortex ($P = 0.002$). These negative correlations remained significant with the application of global normalization in the reanalysis ($P < 0.001$).

Discussion

These combined fMRI and SRH results provide new insights into the brain mechanisms of reality distortion during AVH related to schizophrenia or schizoaffective disorder. First, among the widespread neuronal network of AVH-related activations, the signals from the bilateral IFG correlated strongly with the SRH, independently of the subjective interpretation style, loudness of AVHs and the dose of antipsychotic medication.

The motor theory of speech perception assumes the comprehension of external speech to rely on subliminal matching to the listener's own articulatory gestures (Lieberman and Whalen, 2000). This matching probably relies on the IFG pre-motor speech-production area that is activated also during speech comprehension; for a review, see Nishitani *et al.* (2005). Although imaging studies of AVHs have focused on auditory cortices (Allen *et al.*, 2008), our findings converge with theoretical literature (Atkinson, 2006) and with recent imaging findings with a larger subject group (Sommer *et al.*, 2008) to suggest that the IFG correlates of the SRH comprise the perceptual key substrate for AVHs.

In addition to speech comprehension, the IFG is involved in the production of overt and inner speech, as well as in the imagination of the speech of others (McGuire *et al.*, 1996; Lieberman and Whalen, 2000; Nishitani *et al.*, 2005). Therefore, additional brain circuits are likely to contribute to differentiation between self-produced and externally triggered verbal material (Jeannerod and Pacherie, 2004). We expected neuronal substrates for this distinction to include cortical midline regions related to self-monitoring (Northoff and Berman, 2004). Accordingly, the SRH correlated negatively with the AVH-related coupling of the IFG with the pregenual and posterior cingulate cortex. The decreased coupling could reflect the poor controllability of AVHs that is likely to be associated with the SRH. This interpretation agrees with the finding that the function of the medial prefrontal cortex near the pregenual cingulate cortex is compromised when subjects with schizophrenia attribute words as self-generated or as generated by the experimenter (Vinogradov *et al.*, 2008).

Our analysis of the coupling of the IFG aimed further to test the hypotheses about association of the SRH with the agency-related and the salience-detection-related circuitries that have been suggested to contribute to reality distortion (Kapur, 2003; Frith, 2005). Supporting the first hypothesis, SRH scores correlated positively with the coupling of the IFG with the middle anterior cingulate cortex and the right posterior temporal lobe. The middle anterior cingulate cortex has been implicated in the experience of agency in healthy subjects (Tomlin *et al.*, 2006) and in misattribution of own recorded speech as others' speech in schizophrenic subjects (Allen *et al.*, 2007) and the right posterior temporal lobe has been implicated in the experience of agency (Tankersley *et al.*, 2007).

During treatment with antipsychotic drugs, the reduction in reality-distortion-related positive symptoms correlates with the

strength of the drug's dopamine-2-receptor (D2) binding in the striatum (Seeman *et al.*, 1976; Agid *et al.*, 2007). The striatum is involved in neuronal correlates of salience and aberrant salience of environmental features and internal representations may relate to both delusions and hallucinations (Kapur, 2003). Thus, the coupling of striatum with the IFG could reflect subjective salience of AVH and therefore the SRH.

The observed coupling of the IFG with the auditory cortex is in line with earlier findings of changes in the structural connections between the IFG and the temporal lobe in subjects experiencing AVH (Hubl *et al.*, 2004). The positive correlation of the SRH with this coupling probably relates to auditory features of AVH, in agreement with the observed correlation of the SRH with the subjective loudness of AVH.

In conclusion, our findings are the first to demonstrate how brain activation and coupling within a large-scale neuronal network during a reality distortion symptom (in this case AVH) relate to the subjective reality of the symptom. Because subjects' symptom scores were low (except AVH), further studies are needed to resolve whether similar brain correlates of the SRH occur during exacerbation periods and acute psychosis. Whether similar circuitries are related to other forms of reality distortion also remains to be studied. It is likely that the IFG and auditory cortex are related to the specific form of hallucination that we studied (AVH), whereas other circuitries, whose coupling with the IFG correlated with the SRH, could be involved more generally in reality distortion.

Acknowledgements

We thank Jaana Hiltunen for offering her advice during the analysis, Veikko Jousmäki and Helge Kainulainen for establishing the rating system, Marita Kattelus for her assistance in the MR scanning, Sanna Malinen (auditory stimuli) and Antti Tarkiainen for writing the Presentation[®] script to run the stimuli, Marko Manninen for conducting the neuropsychological testing, Mika Seppä for helpful comments and Cathy Nangini and Stephen Stalter for editing the manuscript. And our heartfelt thanks go to the voice-hearing participants!

Funding

This study benefited from the financial support of the Academy of Finland (National Centres of Excellence Programme 2006–2011), ERC Advanced Grant #232946, the Ella and Georg Ehrnrooth Foundation, the Finnish Medical Foundation, the Jalmar and Rauha Ahokas Foundation, and the Finnish Psychiatric Research Foundation.

Supplementary material

Supplementary material is available at *Brain* online.

References

Agid O, Mamo D, Ginovart N, Vitcu I, Wilson AA, Zipursky RB, *et al.* Striatal vs extrastriatal dopamine D2 receptors in antipsychotic

response—a double-blind PET study in schizophrenia. *Neuropsychopharmacology* 2007; 32: 1209–15.

Allen P, Amaro E, Fu CH, Williams SC, Brammer MJ, Johns LC, *et al.* Neural correlates of the misattribution of speech in schizophrenia. *Br J Psychiatry* 2007; 190: 162–9.

Allen P, Laroi F, McGuire PK, Aleman A. The hallucinating brain: A review of structural and functional neuroimaging studies of hallucinations. *Neurosci Biobehav Rev* 2008; 32: 175–91.

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th edn., Washington, DC: American Psychiatric Press; 1994.

Andreasen NC. Linking mind and brain in the study of mental illnesses: a project for a scientific psychopathology. *Science* 1997; 275: 1586–93.

Atkinson JR. The perceptual characteristics of voice-hallucinations in deaf people: insights into the nature of subvocal thought and sensory feedback loops. *Schizophr Bull* 2006; 32: 701–8.

Beck AT, Baruch E, Balter JM, Steer RA, Warman DM. A new instrument for measuring insight: the Beck Cognitive Insight Scale. *Schizophr Res* 2004; 68: 319–29.

Bingel U, Rose M, Glascher J, Büchel C. fMRI reveals how pain modulates visual object processing in the ventral visual stream. *Neuron* 2007; 55: 157–67.

Bullmore E, Brammer M, Williams SC, Rabe-Hesketh S, Janot N, David A, *et al.* Statistical methods of estimation and inference for functional MR image analysis. *Magn Reson Med* 1996; 35: 261–77.

Centorrino F, Eakin M, Bahk WM, Kelleher JP, Goren J, Salvatore P, *et al.* Inpatient antipsychotic drug use in 1998, 1993, and 1989. *Am J Psychiatry* 2002; 159: 1932–5.

David AS. Auditory hallucinations: phenomenology, neuropsychology and neuroimaging update. *Acta Psychiatr Scand Suppl* 1999; 395: 95–104.

Dierks T, Linden DE, Jandl M, Formisano E, Goebel R, Lanfermann H, *et al.* Activation of Heschl's gyrus during auditory hallucinations. *Neuron* 1999; 22: 615–21.

Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K, *et al.* A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage* 2005; 25: 1325–35.

Friston KJ. Dysfunctional connectivity in schizophrenia. *World Psychiatry* 2002; 1: 66–71.

Friston KJ, Büchel C, Fink GR, Morris J, Rolls E, Dolan RJ. Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage* 1997; 6: 218–29.

Frith C. The self in action: lessons from delusions of control. *Conscious Cogn* 2005; 14: 752–70.

Frith C, Done J. Positive symptoms of schizophrenia. *Br J Psychiatry* 1989; 154: 569–70.

Frith CD. The cognitive neuropsychology of schizophrenia. Hove: Erlbaum; 1992.

Gitelman DR, Penny WD, Ashburner J, Friston KJ. Modeling regional and psychophysiological interactions in fMRI: the importance of hemodynamic deconvolution. *Neuroimage* 2003; 19: 200–7.

Hoffman RE, Hampson M, Wu K, Anderson AW, Gore JC, Buchanan RJ, *et al.* Probing the pathophysiology of auditory/verbal hallucinations by combining functional magnetic resonance imaging and transcranial magnetic stimulation. *Cereb Cortex* 2007; 17: 2733–43.

Hubl D, Koenig T, Strik W, Federspiel A, Kreis R, Boesch C, *et al.* Pathways that make voices: white matter changes in auditory hallucinations. *Arch Gen Psychiatry* 2004; 61: 658–68.

Hunter MD. Locating voices in space: a perceptual model for auditory hallucinations? *Cognit Neuropsychiatry* 2004; 9: 93–105.

Hustig HH, Hafner RJ. Persistent auditory hallucinations and their relationship to delusions and mood. *J Nerv Ment Dis* 1990; 178: 264–7.

Jeannerod M, Pacherie E. Agency, simulation and self-identification. *Mind Lang* 2004; 19: 113–46.

Kaplan RD, Szechtman H, Franco S, Szechtman B, Nahmias C, Garnett ES, *et al.* Three clinical syndromes of schizophrenia in untreated subjects: relation to brain glucose activity measured by positron emission tomography (PET). *Schizophr Res* 1993; 11: 47–54.

- Kapur S. Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. *Am J Psychiatry* 2003; 160: 13–23.
- Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull* 1987; 13: 261–76.
- Kriegeskorte N, Simmons WK, Bellgowan PSF. Circular analysis in systems neuroscience: the dangers of double dipping. *Nat Neurosci* 2009; 12: 535–40.
- Lennox BR, Park SB, Medley I, Morris PG, Jones PB. The functional anatomy of auditory hallucinations in schizophrenia. *Psychiatry Res* 2000; 100: 13–20.
- Lieberman AM, Whalen DH. On the relation of speech to language. *Trends Cogn Sci* 2000; 4: 187–96.
- Liddle PF, Friston KJ, Frith CD, Frackowiak RS. Cerebral blood flow and mental processes in schizophrenia. *J R Soc Med* 1992; 85: 224–7.
- Macaluso E, Frith CD, Driver J. Modulation of human visual cortex by crossmodal spatial attention. *Science* 2000; 289: 1206–8.
- McGuire PK, Silbersweig DA, Murray RM, David AS, Frackowiak RS, Frith CD. Functional anatomy of inner speech and auditory verbal imagery. *Psychol Med* 1996; 26: 29–38.
- Nasar S. *A beautiful mind*. New York: Touchstone Book; 1998.
- Nayani TH, David AS. The auditory hallucination: a phenomenological survey. *Psychol Med* 1996; 26: 177–89.
- Nishitani N, Schürmann M, Amunts K, Hari R. Broca's region: from action to language. *Physiology (Bethesda)* 2005; 20: 60–9.
- Northoff G, Bermpohl F. Cortical midline structures and the self. *Trends Cogn Sci* 2004; 8: 102–7.
- Oulis PG, Mavreas VG, Mamounas JM, Stefanis CN. Clinical characteristics of auditory hallucinations. *Acta Psychiatr Scand* 1995; 92: 97–102.
- Pasley BN, Mayes LC, Schultz RT. Subcortical discrimination of unperceived objects during binocular rivalry. *Neuron* 2004; 42: 163–72.
- Seeman P, Lee T, Chau-Wong M, Wong K. Antipsychotic drug doses and neuroleptic/dopamine receptors. *Nature* 1976; 261: 717–9.
- Shergill SS, Brammer MJ, Williams SC, Murray RM, McGuire PK. Mapping auditory hallucinations in schizophrenia using functional magnetic resonance imaging. *Arch Gen Psychiatry* 2000; 57: 1033–8.
- Silbersweig DA, Stern E, Frith C, Cahill C, Holmes A, Grootenck S, et al. A functional neuroanatomy of hallucinations in schizophrenia. *Nature* 1995; 378: 176–9.
- Sommer IE, Diederer KM, Blom JD, Willems A, Kushan L, Slotema K, et al. Auditory verbal hallucinations predominantly activate the right inferior frontal area. *Brain* 2008; 131: 3169–77.
- Stephan KE, Marshall JC, Friston KJ, Rowe JB, Ritzl A, Zilles K, et al. Lateralized cognitive processes and lateralized task control in the human brain. *Science* 2003; 301: 384–6.
- Stephane M, Thuras P, Nasrallah H, Georgopoulos AP. The internal structure of the phenomenology of auditory verbal hallucinations. *Schizophr Res* 2003; 61: 185–93.
- Surguladze S, Russell T, Kucharska-Pietura K, Travis MJ, Giampietro V, David AS, et al. A reversal of the normal pattern of parahippocampal response to neutral and fearful faces is associated with reality distortion in schizophrenia. *Biol Psychiatry* 2006; 60: 423–31.
- Tabachnick B, Fidell L. *Using multivariate statistics*. 4th edn., Needham Heights: Allyn & Bacon; 2001. p. 145–7.
- Tankersley D, Stowe CJ, Huettel SA. Altruism is associated with an increased neural response to agency. *Nat Neurosci* 2007; 10: 150–1.
- Taylor SF, Welsh RC, Chen AC, Velander AJ, Liberzon I. Medial frontal hyperactivity in reality distortion. *Biol Psychiatry* 2007; 61: 1171–8.
- Tiihonen J, Hari R, Naukkarinen H, Rimon R, Jousmäki V, Kajola M. Modified activity of the human auditory cortex during auditory hallucinations. *Am J Psychiatry* 1992; 149: 255–7.
- Tomlin D, Kayali MA, King-Casas B, Anen C, Camerer CF, Quartz SR, et al. Agent-specific responses in the cingulate cortex during economic exchanges. *Science* 2006; 312: 1047–50.
- van de Ven VG, Formisano E, Roder CH, Prvulovic D, Bittner RA, Dietz MG, et al. The spatiotemporal pattern of auditory cortical responses during verbal hallucinations. *Neuroimage* 2005; 27: 644–55.
- Wechsler D. *The Wechsler Adult Intelligence Scale-Revised*. New York: The Psychological Corporation; 1981.
- Vinogradov S, Luks TL, Schulman BJ, Simpson GV. Deficit in a neural correlate of reality monitoring in schizophrenia patients. *Cereb Cortex* 2008; 18: 2532–9.
- Woodruff PW. Auditory hallucinations: insights and questions from neuroimaging. *Cognit Neuropsychiatry* 2004; 9: 73–91.