NeuroImage: Clinical 20 (2018) 830–843

Contents lists available at ScienceDirect

NeuroImage: Clinical

journal homepage: www.elsevier.com/locate/ynicl

Draining the pond and catching the fish: Uncovering the ecosystem of auditory verbal hallucinations

Jasper Looijestijn[a,b,*], Jan Dirk Blom[a,c,d], Hans W. Hoek[a,c,e], Remco Renken[b], Edith Liemburg[b,c], Iris E.C. Sommer[c,f], André Aleman[b], Rutger Goekoop[a]

[a] Parnassia Psychiatric Institute, Kiwistraat 43, 2552 DH The Hague, the Netherlands
[b] Neuroimaging Center, University Medical Center Groningen, University of Groningen, Antonius Deusinglaan 2, 9713 AW Groningen, the Netherlands
[c] Department of Psychiatry, University Medical Center Groningen, University of Groningen, Hanzeplein 1, 9700 RB Groningen, the Netherlands
[d] Faculty of Social and Behavioural Sciences, Leiden University, Wassenaarseweg 52, 2333 AK Leiden, the Netherlands
[e] Department of Epidemiology, Columbia University, 722 West 168th St, New York, NY, USA
[f] Psychiatry Department, University Medical Center Utrecht & Rudolf Magnus Institute for Neuroscience, Utrecht, the Netherlands

ARTICLE INFO

Keywords:
Schizophrenia
Functional MRI
Graph analysis
Effective connectivity
Systems biology

ABSTRACT

The various models proposed for the mediation of auditory verbal hallucinations (AVH) implicate a considerable number of brain areas and mechanisms. To establish which of those mechanisms are actually involved in the mediation of AVH, we developed a novel method to analyze functional MRI data, which allows for the detection of the full network of mutually interacting brain states, and the identification of those states that are relevant to the mediation of AVH, while applying a minimum number of preconceived assumptions. This method is comparable to the draining of a pond to lay bare the full ecosystem that affects the presence of a particular fish species. We used this model to analyze the fMRI data of 85 psychotic patients experiencing AVH. The data were decomposed into 98 independent components (ICs) representing all major functions active in the brain during scanning. ICs involved in mediating AVH were identified by associating their time series with the hallucination time series as provided by subjects within the scanner. Using graph theory, a network of interacting ICs was created, which was clustered into IC modules. We used causal reasoning software to determine the direction of links in this network, and discover the chain of events that leads to the conscious experience of hallucinations. Hallucinatory activity was linked to three of the seven IC clusters and 11 of the 98 ICs. ICs with the most influential roles in producing AVH-related activity were those within the so-called salience network (comprising the anterior cingulate gyrus, right insula, Broca's homologue, premotor cortex, and supramarginal gyrus). Broca's area and the cerebellar regions were significantly, but more distantly involved in the mediation of AVH. These results support the notion that AVH are largely mediated by the salience network. We therefore propose that the mediation of AVH in the context of schizophrenia spectrum disorders involves the attribution of an excess of negative salience by anterior-cingulate areas to linguistic input from Broca's right homologue, followed by subsequent processing errors in areas further 'downstream' the causal chain of events. We provide a detailed account of the origin of AVH for this patient group, and make suggestions for selective interventions directed at the most relevant brain areas.

Non-standard abbreviations

SM module	sensorimotor module
C-E-R module	cognition evaluation response module
VI-EM module	visual imagery/episodic memory module

1. Introduction

Auditory verbal hallucinations (AVH) are the most prevalent types of hallucination in individuals diagnosed with a schizophrenia spectrum disorder, as well as in individuals without a diagnosis, psychiatric or otherwise (Ohayon, 2000). They have been the object of extensive neuroimaging research over the last 20 years and various hypotheses have been proposed concerning their mediation (Blom, 2015). As

* Corresponding author.
E-mail address: j.looijestijn@parnassia.nl (J. Looijestijn).

https://doi.org/10.1016/j.nicl.2018.09.016
Received 15 June 2018; Received in revised form 9 September 2018; Accepted 18 September 2018
Available online 21 September 2018

recently summarized by Curcic-Blake et al. (Curcic-Blake et al., 2017), the four major hypotheses involve: i) memory intrusion into language processing, ii) disrupted self-monitoring of inner speech, iii) aberrant cerebral lateralization, and iv) unbalanced top-down and bottom-up processing. As all four models overlap somewhat with respect to the brain regions involved, hybrid models for AVH postulate an overflow of default-mode-network-derived information into sensory association cortices or into central executive networks (CEN), with the ensuing intrusions of these networks being falsely attributed to an external source (Northoff & Qin, 2010; Jardri et al., 2013; Palaniyappan, 2012; Looijestijn et al., 2015). In the 'triple network model' it is hypothesized that such imbalances between the default-mode network (DMN) and CEN are mediated by a disrupted function of the salience network (SN) when the latter fails to attribute appropriate salience to input from resting-state and active-state (central executive) modi. The triple network model has been studied by Manoliu et al. (2013), where they find decreased activity of the right anterior insula to be associated with increased connectivity between the DMN and CEN, as well as hallucination severity. Alternatively, intrusions into CENs are thought to derive from unstable neural networks at a lower spatiotemporal level that erratically switch between their high-frequency active state and their low-frequency resting state (Looijestijn et al., 2015; Loh et al., 2007).

Functional MRI studies have reported that hallucination-related brain activity precedes the conscious experience of hallucinations by as much as nine seconds, which is way before subjects become conscious of the hallucination (Diederen et al., 2010; Hoffman et al., 2011; Shergill, 2004). The experience of hallucinations therefore is thought to depend on a chain of neural events that precedes it. In a recent study by Lefebvre et al. (2016), examination of the triple network model is elaborated by including the hippocampus in their hypothesis and they focus on transitions between different phases surrounding the hallucinatory state. Evidence was found that increased hippocampal input to the SN is involved in switching to a hallucinatory state. Notwithstanding, the nature of the causal chain of events has so far remained largely unclear, since current methods of functional imaging were limited in their ability to examine this chain of events. Most fMRI studies have used strongly model-based methods to identify brain activity, which involves searching the brain for specific patterns of interest. As the detection of hallucinations in fMRI studies is especially challenging due to their unprompted occurrence in subjects (Leroy et al., 2017), this often requires strongly specified models to obtain statistical power; or, in the case of data-driven studies, a selective use of data (components of interest) (Jardri et al., 2013; Manoliu et al., 2013; Lefebvre et al., 2016; van de Ven et al., 2005). Such approaches can be compared to fishing with a matched spinner for one particular type of fish, which is nonetheless part of a complex ecosystem. Thus, model-based methods are confirmatory methods, which provide information on expected patterns, but these should be complemented by exploratory methods that allow for the discovery of unexpected (yet relevant) findings. In the fishing analogy, we would ideally want to employ a method that allows us to drain the pond without losing important species (e.g. noise reduction) to uncover the entire ecosystem (all neural events within the brain), after which we can select all species (neural events) that affect the presence of our main fish of interest (i.e. neural activity that is directly related to the conscious experience of AVH). In this paper, we present such a method, and use it to discover the chain of events that lead up to the conscious experience of AVH. The clinical relevance of these findings is shortly discussed.

Our method involves the use of a so called 'model-free analysis' of functional brain connectivity concomitant with AVH, based on fMRI data obtained from 85 hallucinating patients who were diagnosed with a schizophrenia spectrum disorder. Bayesian network analyses (Mumford & Ramsey, 2014) allowed to test assumptions regarding the direction of the causal influence of implicated brain regions on each other. Using a minimum of a priori assumptions about the nature of event-related brain activity, we provide i) a systematic account of the circuits mediating AVH in the patient group, and ii) a perspective on the mediation of AVH that is complementary to that of model-based studies. The challenge here is to *search*, among the vast number of available hypotheses, for the hypothesis that explains the data best, and - preferentially - also facilitates therapeutic interventions.

2. Materials and methods

2.1. Participants

A total of 85 right-handed patients experiencing frequent AVH (i.e., at least three episodes per 15 min) were recruited at Parnassia Psychiatric Institute and the University Medical Center Utrecht. Exclusion criteria included the presence of neurological disorders, IQ < 80, structural brain deficits, and coarse scanner artefacts upon initial inspection of the fMRI data. Of all patients, 56% were male; mean age was 38 (SD 11.0) years, and average time spent on education was 12.5 (SD 2.5) years. All patients were diagnosed in accordance with the DSM-IV-TR criteria as suffering from Schizophrenia (77%), Schizoaffective Disorder (3%) or Psychotic Disorder Not Otherwise Specified (20%). Diagnostic interviews had been carried out by independent psychiatrists using the Comprehensive Assessment of Symptoms and History (CASH) (Andreasen et al., 1992). There was a large range in the number of years since the onset of hallucinations, with a mean duration of 14.5 (SD 12.5) years. Total PANSS-score averaged 68 (SD 15.5). The majority of participants used antipsychotic medication (89%), with a mean chlorpromazine-equivalent dose of 413 (SD 318) mg/d (Woods, 2003). Of the medicated participants, 36% used clozapine, 34% other second-generation antipsychotics, 26% first-generation antipsychotics, and 4% a combination of these. As these figures indicate, a substantial number of our subjects was treatment-refractory with respect to their AVH. After the participants had received a complete description of the study, written informed consent was obtained in accordance with the Declaration of Helsinki. The study was approved by the Human Ethics Committee of the University Medical Center Utrecht. Looijestijn et al. (2013) previously reported on a subset of the fMRI data of these patients (52 of the 85 subjects), presenting the results of a model-based analysis of AVH perceived inside the head (internal AVH) versus those perceived as coming from outside the head (external AVH).

2.2. Image acquisition

Functional neuroimaging maps were obtained with a Philips Achieva 3 Tesla Clinical MRI scanner using a fast 3D PRESTO SENSE sequence, achieving full brain coverage within 0.609 s (Neggers et al., 2008). PRESTO (PRinciple of Echo Shifting with a Train of Observations) makes optimal use of the time lapse between excitation by the radiofrequency pulse and readout, by applying the next excitation well before signal readout. The acquisition speed was further enhanced by combining PRESTO with parallel imaging techniques (sensitivity encoding; SENSE), thus allowing for a readout of fewer lines in K-space (Pruessmann et al., 1999). Scanning resulted in 800 3D images, depicting BOLD contrast acquired at the following parameter settings: 40 coronal slices, TR/TE 21.75/32.4 ms per slice, flip angle 10°, FOV 224 × 256 × 160 mm, matrix 64 × 64 × 40, voxel size 4 mm isotropic. The total functional imaging time per patient was approximately 8 min and 7 s. During the scanning sessions, participants were instructed to squeeze a balloon whenever they experienced AVH and to release it when the hallucinations subsided. A high-resolution anatomical scan with parameters TR/TE 9.86/4.6 ms, 1 × 1 × 1 mm voxel size, flip angle 8°, was acquired to improve localisation of the functional data.

2.3. Preprocessing

The FMRIB software library (FSL, Oxford, http://www.fmrib.ox.ac.

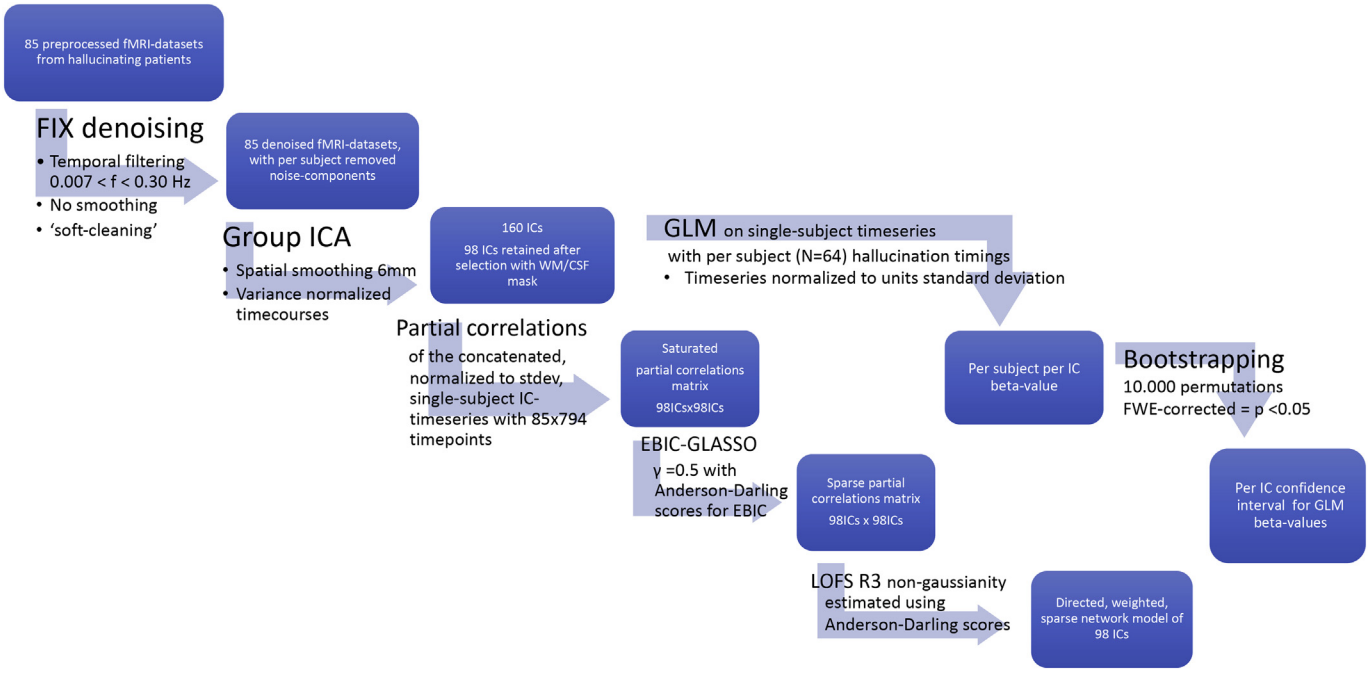

Fig. 1. Analysis pipeline. Used methods in an analysis pipeline, with output in blue boxes and directly below or next to the boxes the processing steps before input into the next phase of analysis.

uk/fsl/) was used for data analysis. Prestatistical processing consisted of motion correction (Jenkinson et al., 2002) and non-brain tissue removal. Six initial volumes were deleted to reach steady-state imaging. Temporal band-pass filtering was applied, using a liberal bandwidth ($0.007 < f < 0.30$ Hz) to maintain a broad range of frequencies, thus allowing for possible high-frequency AVH-related brain activity and upholding non-gaussianity in the data to perform causal searches (Ramsey et al., 2014). This broad temporal range allowed us to delineate a greater number of (subdivided) functional networks (Smith et al., 2011). The 0.30 Hz cut-off was chosen to thoroughly remove a scanner artefact settled around 0.38 Hz. Individual fMRI data were denoised in three steps using the novel FMRIB's ICA-based Xnoisefier (FIX), a data-driven automated classifier of signal-versus-noise components (Salimi-Khorshidi et al., 2014; Griffanti et al., 2014). No spatial smoothing was applied at initial pre-processing, as this might have reduced the sensitivity of ICA (and hence FIX) to certain artefacts and signals (Salimi-Khorshidi et al., 2014). Fig. 1 shows an analysis pipeline for an overview of the methods we used.

2.4. Denoising

Since all fMRI data is inherently noisy (due to spurious signals produced by movement, respiratory motion, pulsation, and MRI acquisition) we started out with a thorough denoising process (Van Dijk et al., 2012). The first step of this process involved *training*. To optimize FIX for the fMRI PRESTO task, we used a subset consisting of the first 33 participants (recruited from an alphabetically arranged list) for hand-training of the classifier. Thus subject-level independent components (IC) from the independent-component analysis (ICA) in FSL (Beckmann & Smith, 2004) were assessed with regard to temporal and spatial characteristics by two raters from our study group and one external rater, all of whom scored the results either as 'signal' or 'noise'. Spatial maps were assessed for noise from i) cardiac pulsation, ii) movement, iii) susceptibility artefacts, iv) white matter fluctuations, v) the sagittal sinus, and vi) MRI acquisition. The intra-class correlation coefficient after the first (independent) pass of our three raters was 0.71 ($F = 3.4$, $df = 1993$, $p = 0$) and Fleiss kappa was 0.36 ($Z = 27.5$, $p = 0$), both signifying reasonable agreement. A second non-independent pass (consensus meeting) was performed to reach a unanimous decision by all three raters. Any ambiguous components were given the benefit of the doubt in order to prevent the loss of valuable information. The ensuing classifications were fed into FIX to train the multi-level classifier. The second step involved *classification*. During this stage, the resulting training file was used by the FIX algorithm to classify the ICs of all 85 participants as 'signal' or 'noise'. FIX requires a threshold for classification to be chosen (of 1–100) for the level of signal-versus-noise components. We used a classification threshold of 40, based on the highest true-positive and false-negative rating results of the Leave One Out-testing (LOO-testing) (Salimi-Khorshidi et al., 2014), and confirmed these by manually inspecting all signal-versus-noise classification decisions. The third and final step involved *cleanup*, meaning that all noise components were subtracted from the individual fMRI datasets ('*soft-cleaning*') (Griffanti et al., 2014), including motion confounders, yielding 85 preprocessed and denoised fMRI datasets for further analysis.

2.5. Group-level independent component analysis: identification of ICs

During the next stage, we used a second run of GICA with automatic component estimation in FSL (Beckmann & Smith, 2004). At this phase normalization of the voxel-wise variance was done to focus the estimation on temporal dynamics and have less influence from a voxel's mean signal. The preprocessed functional data, containing 794 time points for each participant, were temporally concatenated across patients to create a single 4D data set. The resulting 160 ICs were visually inspected to identify any remaining artefacts using a white-matter/cerebrospinal-fluid mask (WM/CSF mask), based on averaged individual anatomical scans. Whenever the local maxima of IC spatial maps were located inside the WM/CSF mask (or whenever the IC constituted a clear rim artefact), group-level ICs were excluded from further analysis. If there were any doubts regarding the nature of the signal, ICs were not excluded (IC2, IC68). As a result, of the initial 160 ICs, 98 were retained for further analysis.

2.6. Constructing a sparse directed IC network

The following stage involved the construction of a multimodular, directed IC network that would allow to estimate the effective connectivity (e.g., the causal directions) of the various links between ICs. As Dynamic Causal Modelling and Granger Causality are highly controversial for use in fMRI (Friston et al., 2003; David et al., 2008, Seth et al., 2015), we opted for Bayesian network-modelling techniques. Most problems regarding the inference of causal directions in fMRI data can be overcome using these techniques (Ramsey et al., 2010), which have been tested on simulated fMRI data (Smith et al., 2011; Ramsey et al., 2010; Ramsey et al., 2011) showing ≥95% accuracy (Ramsey et al., 2014). Essential to this approach is to i) start by applying a model-selection algorithm to reduce the number of links, then ii) create an undirected sparse graph, and, during the next stage, iii) use non-gaussian information in the skeleton graph to estimate causal directions.

First, to establish links between the various ICs, single-subject time courses were reconstructed by regressing group-spatial maps into each subject's 4D dataset (Filippini et al., 2009). Next, the time courses of the 98 individual ICs were concatenated (98 ICs with 85 × 794 time points) to calculate group-level covariance matrices. This yielded a fully saturated network with 98 × 98 links, even though some correlations were weak. Secondly, we used EBIC-glasso (Extended Bayesian Information Criterion, graphical least absolute shrinkage and selection operator) (Foygel & Drton, 2010; van Borkulo et al., 2014), as implemented in the R-package *qgraph* (psychosystems.org) (Epskamp et al., 2012), to perform initial model selection. EBIC-glasso is a data-driven method that employs a measure of information conservation (the EBIC) (Chen & Chen, 2012) to optimally converge onto a network solution that possesses a high sparsity, but still succeeds in properly explaining the data. Glasso (Friedman et al., 2008) is a regularization technique for fast estimations of optimal models in large networks. The basis for these estimations is a saturated partial correlation matrix where spurious connections are controlled for by means of a tuning parameter λ (lambda) for the penalization of the maximum likelihood estimation. It thus creates 100 network solutions, ranging from fully saturated to fully disconnected. From this range of networks, the graph with an optimal solution of sparsity while still representing the data (i.e., the EBIC score) was selected. Covariance-based methods using regularization techniques are accurate in estimating the presence of network connections across a range of fMRI conditions (Smith et al., 2011). EBIC scores have been used successfully in fMRI studies that aimed to obtain sparse network models (Manelis et al., 2016; Dobryakova et al., 2015) while investigating limited sets of nodes (e.g., regions of interest, ROIs). The hyperparameter γ (gamma) was set to a default of 0.5, which produces optimal solutions in most simulated datasets (Foygel & Drton, 2010). Third, the Linear Non-Gaussian Orientation, Fixed Structure (LOFS) algorithm was used to estimate the direction of links with the aid of the R3 rule (Ramsey et al., 2014). LOFS uses rules that (like the LiNGAM algorithm) infer orientation in a linear, non-Gaussian system, while orienting links in a pairwise manner, without reference to the additional context in the graph. Effective connectivity is established by estimating the model with the highest non-gaussianity of the error term. We estimated the degree of non-gaussianity by using Anderson-Darling scores (Anderson & Darling, 1952). The R3 rule was chosen for its conservative character in appointing causal links in combination with high accuracy (Ramsey et al., 2014). Links between functional networks are expected to be reciprocal and, as such, we did not want to force direction, and only attain dominant directions of influence. Links with ambiguous directions were conceptually taken as bidirectional. Graph analyses were conducted using TETRAD-V (v.5.3.0; http://www.phil.cmu.edu/projects/tetrad).

2.7. Modularity

We detected a modular structure in the sparse network by using the Louvain algorithm developed by Blondel et al. (Blondel, 2008), a search algorithm that optimizes modularity, which is capable of using weighted links and detecting nested clusters (i.e., smaller clusters within larger clusters). A methodological study by Rubinov and Sporns (Rubinov & Sporns, 2011) concluded that modularity primarily depends on the relative difference between weight magnitudes. Therefore, in the analysis we decided to include absolutes of negative links (7%), as these can provide valuable extra information regarding the organization of networks by indicating instances of *deactivation* of functionally related ICs. In conformity with the calculations of links in the IC network, we also detected interactions at the level of IC modules, and thus created an IC-network graph.

2.8. General linear model with balloon presses

The function of the various ICs and IC modules within the multimodular IC network was inferred by i) examining their spatial patterns, and comparing them with previous reports on the function of such networks, and ii) linking hallucinations to individual IC-time series with the aid of within-scanner hallucination timings. The time courses of consciously experienced AVH (as indicated by individual patients within the scanner with the aid of balloon presses) were linked to the time courses of subject-level ICs by performing a post-hoc general linear model (GLM) analysis with subject-level IC time series as an independent (to be explained) variable and the subject-level model of the BOLD response to the AVH as a dependent (explanatory) variable. The BOLD responses coinciding with AVH were modeled by a boxcar based on the within-scanner balloon presses, which was subsequently convolved with a single gamma function without post-stimulus undershoot to model the hemodynamic response. Excluded from this part of the analysis were patients who exhibited continuous AVH (n = 2), recorded no AVH during scanning (n = 1), reported difficulties with the balloon presses (n = 3), or showed > 50% ambiguous (n = 10) or missing (n = 5) AVH responses. Consequently, 64 patients qualified for this part of the analysis. The average duration of hallucinations in the scanner was 17.7 s (SD 47.3), with a large range of 0.8–61.5 s across subjects. Likewise, the frequency of AVH ranged from 1 to 129 hallucinations per scanning session (average 20, SD 21). Each participant's AVH model was tested with the 98 ICs remaining after discarding noise ICs, resulting in 64 participants × 98 beta values. Calculated beta values were fed into a bootstrapping procedure (n = 10,000 repetitions) in Matlab to create an across-ICs confidence interval. We used an FWE threshold at $p < 0.05$ to identify AVH-related ICs, and also took the numerical beta value into account to keep an open perspective and value relative activation and deactivation of ICs during AVH. These post-hoc analyses provided information on the 'distance' of ICs to hallucinations and, thus, yielded information on the positioning of hallucination-related ICs in the IC network as a whole.

2.9. Network metrics

Individual ICs were examined for the singular influence that they were likely to have in directing the flow of information through the network of IC correlations. Interconnecting hubs were calculated using the betweenness-centrality measure adapted for weighted networks, with high values indicating that ICs participated in a large number of relatively short paths between the ICs of the network (Opsahl et al., 2010). As interconnecting hubs comprise a large part of the information flow through networks (often bridging different modules), they are considered crucial for efficient communication and control of networks (Liu et al., 2011). Furthermore, we calculated the weighted degree for a measure of local influence in the network. To investigate sources of brain activity that are *indirectly* linked to AVH (see Introduction), we

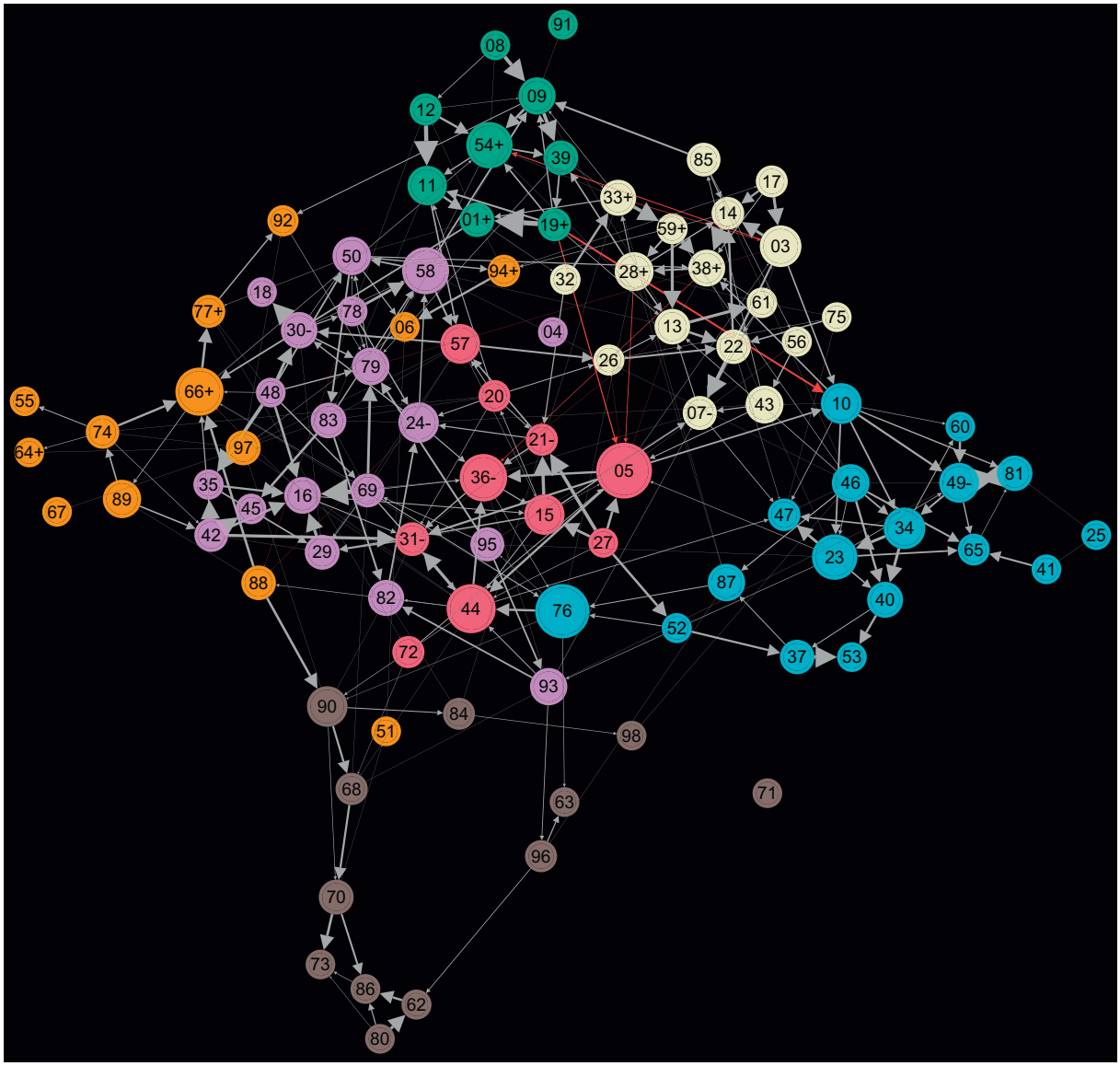

Fig. 2. IC network graph. Network graph visualization of the IC network using the ForceAtlas-algorithm (gephi.org), with edge thickness for partial correlations ($r > 0.02–0.29$) grey color for positive correlations, red color for negative correlations. Node color for modularity (see Table 1). Node size for betweenness centrality. Nodes were automatically assigned coordinates based on a force-directed layout algorithm which treats nodes as positive charges that repulse each other, while being constrained by their links.

calculated the weighted degree using only links with the AVH-related ICs identified in the GLM analysis. The accordingly identified ICs were called 'tributaries' to indicate their hypothesized contributory function in the AVH circuits; a Z-score > 2.56 was used to identify these tributaries and their interconnecting hubs.

3. Results

3.1. Identification of ICs

The multi-modular network of the 98 group-level ICs that we constructed allowed a global view of the 'network context' or 'embedding' of all ICs. Within the IC network, we identified seven modules, together covering 97 of the 98 ICs (for summary slides of all ICs, see Supplementary Materials, Fig. 5). To establish the role of ICs in the flow of information throughout the network, hubs were identified and a network graph was constructed (Fig. 2). In the network graph, nodes were automatically assigned coordinates based on a force-directed layout algorithm that treats nodes as positive charges that repulse each other, while being constrained by their links (Gephi 0.9.1, gephi.org). Permutation tests revealed that, of the 98 ICs, 18 had significant associations with AVH-related balloon presses as recorded within the scanner, of which 11 had positive betas. These 11 ICs, which synchronized with balloon presses, functioned as the anchor points for the interpretation of our data. The ICs that activated or deactivated in relation to AVH clustered together within specific modules. The betweenness-centrality measure indicated ICs with a disproportional influence on information transfer throughout the whole network. Thus, ICs 3, 5, 7, 46, 54, 58, 76, 93, and 97 were identified as *interconnecting hubs*, whereas ICs 9, 11, 13, 14, 39, 50 were identified as tributaries in the AVH-related circuit. Table 1 lists the 98 ICs with their anatomical descriptions, grouped per network module, with beta values for associations with the balloon presses and network metrics. The seven IC modules can be described as follows.

3.2. Module I, the sensorimotor module

Module I contained nine ICs, comprising a number of brain regions

Table 1
Independent components (ICs) per module.

Module	IC	Brain areas	GLM Beta	Betweenness	Tributaries
I Sensorimotor	1	L Precentral, postcentral gyrus (superior), cerebellum	**0.171**	41.6	–
	8	L + R Precentral, postcentral gyrus (inferior)	−0.004	0.0	0.03
	9	L + R STG (anterior to posterior)	0.049	167.7	**0.31**
	11	L + R SPL	0.009	52.7	**0.39**
	12	R Precentral, postcentral gyrus	−0.017	27.1	0.19
	19	L > > R SMG, postcentral gyrus, central opercular cortex	**0.116**	164.7	–
	39	R Insula (posterior), central opercular cortex	0.052	63.7	**0.40**
	54	R SMG, postcentral gyrus	**0.066**	**199.1**	–
	91	L Parietal operculum cortex, STG	0.009	0.7	0.01
II Cognition, evaluation/salience and response formation (C-E-R)	3	L > R Fronto-parietal network	−0.033	**178.3**	0.14
	7	R > L Fronto-parietal network	**−0.064**	**258.9**	0.02
	13	L + R Fronto-parietal-occipital network	0.051	77.7	**0.28**
	14	L IFG (Broca)	0.032	14.4	**0.41**
	17	L MFG	0.009	65.8	0.16
	22	L > > R MFG + IFG + MTG	0.032	54.0	0.00
	26	R SPL + SMG	0.004	22.3	0.07
	28	R > > L Insula (anterior), IFG (Broca)	**0.097**	109.2	0.00
	32	L Postcentral gyrus + precentral gyrus (medial)	0.012	3.2	0.15
	33	L + R SFG (posterior medial, SMA)	**0.094**	120.6	–
	38	L + R SFG (superior medial) + frontal pole + L IFG (Broca) + L + R MTG + R caudate	**0.080**	81.0	–
	43	R MTG (anterior)	−0.011	40.6	0.03
	56	L SMG, angular gyrus, STG (posterior) + MTG	0.009	51.0	0.00
	59	L + R dorsal ACG, paracingulate	**0.067**	34.1	–
	61	R Frontal pole	−0.011	32.7	0.11
	75	R MFG (posterior)	−0.014	7.6	0.00
	85	R Temporal pole, STG anterior	0.047	43.5	0.01
III Cerebellar	6	R Cerebellum (crus)	0.040	50.1	0.16
	51	R Cerebellum (anterior inferior)	0.013	39.3	0.00
	55	L Cerebellum crus	0.020	12.6	0.00
	64	L Cerebellum (medial)	**0.058**	4.1	–
	66	Cerebellum vermis (superior)	**0.092**	165.0	–
	67	L Cerebellum (inferior medial)	−0.010	0.0	0.00
	74	L + R Cerebellum (crus)	−0.002	66.1	0.21
	77	R Cerebellum (medial)	**0.062**	9.8	–
	88	L + R Cerebellum (medial superior)	0.027	56.6	0.17
	89	Cerebellum vermis (inferior)	−0.038	11.2	0.07
	92	R Cerebellum (inferior medial)	0.014	44.9	0.16
	94	R Cerebellum (inferior)	**0.068**	46.6	–
	97	L Cerebellum (inferior medial)	−0.013	**227.6**	0.03
IV Visual imagery/episodic memory (VI-EM)	4	R ITG (posterior)	−0.008	0.0	0.00
	16	L + R Primary visual cortex	−0.024	54.9	0.04
	18	L + R Lateral occipital cortex	−0.018	4.5	0.01
	24	L > R Lateral occipital	**−0.093**	61.6	0.00
	29	L + R Occipital pole, cuneus	−0.037	35.8	0.00
	30	R Lateral occipital (superior), SPL	**−0.07**	107.2	0.01
	35	R + L Occipital pole	0.008	16.2	0.00
	42	R Lingual gyrus	−0.011	56.5	0.09
	45	L > R Lateraal occipital gyrus	−0.021	11.0	0.00
	48	R Occipital fusiform gyrus, lingual gyrus	0.024	16.9	0.10
	50	R > L Cerebellum crus	0.056	135.9	**0.28**
	58	R MTG (temporooccipital), lateral occipital gyrus	−0.019	**235.5**	0.06
	69	R Lingual gyrus	−0.023	103.7	0.05
	78	L > R Temporooccipital fusiform cortex	−0.035	147.6	0.00
	79	L > R Temporal occiptal fusiform cortex	0.008	132.6	0.03
	82	L Lingual gyrus, hippocampus	−0.051	37.8	0.00
	83	L Temporal occipital fusiform cortex	−0.001	114.9	0.02
	93	L + R Hippocampus, parahippocampus	−0.015	**260.5**	0.00
	95	R Hippocampus	−0.026	37.1	0.00

(*continued on next page*)

Table 1 (*continued*)

Module	IC	Brain areas	GLM Beta	Betweenness	Tributaries
V anterior DMN	10	L + R SFG (anterior medial)	−0.050	120.0	0.25
	23	R Frontal pole, paracingulate	−0.055	102.7	0.03
	25	L Frontal orbital cortex	−0.008	53.3	0.00
	34	L > R ACG, L + R frontal orbital cortex and frontal pole	0.014	33.0	0.01
	37	R Thalamus, caudate	−0.006	18.4	0.00
	40	R Caudate	0.012	80.8	0.02
	41	R Frontal orbital cortex	−0.001	0.0	0.00
	46	L + R Putamen	0.029	**232.0**	0.10
	47	Paracingulate R	−0.020	35.7	0.01
	49	L rostral ACG, MFG	**−0.061**	44.3	0.00
	52	R PCG, thalamus	0.009	2.1	0.00
	53	R Thalamus (anterior)	−0.014	3.1	0.00
	60	L Caudate	−0.011	4.2	0.00
	65	R Frontal pole, frontal orbital cortex	−0.006	67.2	0.00
	76	R Thalamus	−0.003	**223.6**	0.08
	81	L Frontal pole	−0.048	32.1	0.00
	87	R > L putamen, pallidum	0.004	136.3	0.06
VI Subcortical	62	R Temporal fusiform cortex, temporale pole	−0.010	1.1	0.00
	63	Brainstem	−0.004	15.8	0.00
	68	R > L Cerebellum (superior anterior)	−0.043	88.3	0.00
	70	Brainstem	0.040	52.1	0.00
	71	L Putamen	0.015	0.0	0.00
	73	Brainstem	0.013	64.7	0.00
	80	L Parahippocampus, hippocampus	0.032	47.8	0.00
	84	L Thalamus	0.038	34.4	0.05
	86	Brainstem	0.000	17.8	0.00
	90	Brainstem + L + R STG	−0.015	152.1	0.00
	96	R Temporal fusiform cortex, parahippocampus	−0.023	93.6	0.00
	98	L Pallidum, amygdala	0.031	99.7	0.00
VII posterior DMN	5	L + R Posterior cingulate, precuneus + L + R lat. Occipital	−0.057	**195.3**	0.16
	15	L + R Precuneus	−0.025	82.5	0.00
	20	R Lateral occipital (superior), SPL	−0.050	47.4	0.00
	21	L + R Precuneus	**−0.064**	64.9	0.00
	27	L + R Posterior cingulate (midcingulate)	−0.027	114.2	0.00
	31	R Precuneus, posterior cingulate	**−0.064**	30.1	0.00
	36	R > L Precuneus	**−0.082**	111.8	0.09
	44	L + R Posterior cingulate, precuneus	−0.041	167.9	0.00
	57	L Lateral occipital cortex superior, R precuneus	−0.035	94.5	0.00
	72	L + R Cerebellum (IX)	−0.034	33.8	0.04
None	2	L ITG, MTG	0.007	0.0	0.00

Brain areas derived from local maxima in Harvard-Oxford brain atlas as implemented in FSL, plus (+) for separated clusters, commas (,) for contiguous activation. Betweenness centrality and tributaries bold for $Z > 2.56$. GLM beta's with bold for $p < 0.05$ (corrected). SMG supramarginal gyrus, STG superior temporal gyrus, MTG middle temporal gyrus, ITG inferior temporal gyrus, SPL superior parietal lobule, SFG superior frontal gyrus, MFG middle frontal gyrus, IFG inferior frontal gyrus, ACG anterior cingulate gyrus.

that we characterize as the *sensorimotor module* (SM module). This module comprises the pre- and postcentral gyri (IC1, IC8, IC54), the supramarginal gyrus (IC19, IC54), frontoparietal opercular cortices (IC 19, IC 39, IC91), the posterior insula (IC 39), and the superior temporal gyrus (STG), which includes Heschl's gyrus and the planum temporale (IC9, IC91). Based on previous reports on the functions of these ICs and brain areas, it is likely that the module as a whole has a central function in auditory and motor processing. Primary and secondary auditory cortices in the posterior STG (IC9) showed no significant activation during AVH.

3.3. Module II, the cognition, evaluation/salience, and response formation (C-E-R) module

Module II contained 17 disparate ICs, representing brain activity in prefrontal regions and brain areas centered around the temporo-parietal junction. Among them, three 'executive' fronto-parietal ICs were discernible, of which two were more lateralized (IC7, IC3), and one more balanced (IC13), which we assume to represent subdivided components of the CEN. The right-sided CEN (IC7) and the left-sided CEN (IC3) appeared to be interconnecting hubs. Additionally, we found several 'cognitive' ICs involving language production (IC28, IC38), working memory, self-referential processing, task coordination (IC38), and motor planning (IC33); of note, these may also be involved in other cognitive functions. This module also contained two ICs (IC28, IC59) that together form the salience network (SN). The SN is involved in risk prediction (i.e., the chance of reward), based on information streams of the highest level of integration (i.e., combined emotional and cognitive information). It continuously weighs the risks that are inherent to any operation (whether involving the self, others or the 'common ground'), potentially resulting in a full change of sensory predictive models, executive functions, and subsequent motor (verbal) actions via the CEN (Pezzulo et al., 2015). We therefore termed this module the *cognition, evaluation/salience, and response formation* module (C-E-R module). The ICs associated with the balloon presses included the right anterior insula and Broca's homologue (IC28), the bilateral supplementary motor area (SMA, IC33), the bilateral frontal pole, the superior frontal gyrus, Broca's area (IC38), and bilateral dorsal anterior cingulate cortex (ACC, IC59). Thus, hallucinatory activity in module II mostly involved cognitive (speech production, self-representation), and evaluative (SN) components, but not the executive parts (CEN) of the C-E-R module. The left-sided CEN mainly bridged the pDMN and the visual-imagery/episodic-memory module (i.e., module IV).

3.4. Module III, the cerebellar module

Module III, the *cerebellar module*, comprised 13 ICs almost exclusively located in different cerebellar regions. Remarkably, our

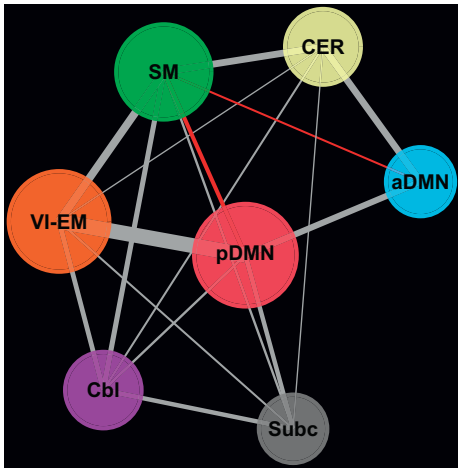

Fig. 3. IC-modules network graph with partial correlations. Edge weight for edge thickness, max partial correlation 0.44, grey color for positive correlations, red color for inverse correlations. Node size for weighted degree. Abbreviations; SM- sensorimotor module, C-E-S – Cognition, evaluation/salience and response formation module, Cb – Cerebellar module, VI-EM– Visual Imagery and Episodic memory module, aDMN- anterior Default Mode Network, pDMN – posterior Default Mode Network, Subc – Subcortical module.

methodological approach revealed an elaborate cerebellar network of more or less separate functional compartments, which is in line with the notion of repeated cerebellar micro-complexes with a different input and output (Apps & Garwicz, 2005), and with limited intracerebellar communication. Four cerebellar ICs (IC64, IC66, IC77, IC94) showed a positive relation with the balloon presses, and might function in the motor control necessary for this activity; or might also reflect the exertion of higher-order cognitive control, i.e., prediction, error monitoring, and online modulation of language and/or speech production (Schmahmann & Sherman, 1998; Andreasen & Pierson, 2008; D'Angelo & Casali, 2012).

3.5. Module IV, the visual-imagery/episodic-memory module (VI/EM module)

The 18 ICs of Module IV comprised mainly occipital brain regions, along with several medial temporal and temporo-occipital regions; therefore, this was called the *visual-imagery/episodic-memory module* (VI/EM module). Three of these ICs (IC21, IC31, IC36) were negatively correlated with the balloon presses. The medial temporal regions, including the hippocampus (IC82, IC93, IC95), showed no significant association with the balloon presses. IC 93, which represents the bilateral hippocampus, appeared as an interconnecting hub joining a network of ICs that deactivated during the AVH (IC7, IC36, IC49) or showed a trend towards deactivation.

3.6. Module V, the anterior default-mode network (DMN) and social-reference module

Module V contained 17 ICs limited to prefrontal regions, the thalamus, and the striatum. We termed it the *anterior DMN and social-reference module* because the medial prefrontal regions of the anterior DMN (Andrews-Hanna et al., 2010; Whitfield-Gabrieli et al., 2011) and the orbitofrontal regions are associated with the integration of limbic areas, the valuation of social cues, and emotion regulation (Bechara et al., 2000; Nieuwenhuis & Takashima, 2011). IC49, which represents the ACC and the medial frontal gyrus, showed a negative association with the balloon presses. Two extrapyramidal regions behaved as interconnecting hubs. The putamen (IC46) mainly bridged the aDMN and C-E-S module, whereas the thalamus (IC76), as expected, was found to function as an interconnecting hub with links throughout all modules.

3.7. Module VI, the subcortical module

Module VI contained 13 ICs representing the brainstem (IC63, IC70, IC73, IC86, IC90), the thalamus, the basal ganglia (IC71, IC84, IC98), the temporal fusiform gyrus, and the parahippocampus (IC62, IC80, IC98). We called this the *subcortical module*. It showed no significant associations with the balloon presses.

3.8. Module VII, the posterior DMN module

Module VII's ICs represented anatomically closely connected regions located in the posterior cingulate and precuneus, with some extensions to lateral visual cortex (IC20, IC57). Its network showed a recognizable similarity to the posterior subdivision of the DMN, and was therefore termed the *posterior DMN module*. Three ICs (IC21, IC31, IC36) deactivated during the AVH, with most of the other ICs showing a trend towards deactivation. The module takes up a central position in the IC network, suggesting that it has considerable influence on information processing throughout the network. The posterior cingulate (IC5) behaved as the module's only interconnecting hub, with a substantial proportion of inverse correlations with extramodular ICs. Interestingly, this module hub positively correlated with the lateralized CEN hubs of the C-E-S module (IC3, IC7), as well as with the module VI hub, the hippocampus (IC 93).

3.9. Effective connectivity

The EBIC-glasso algorithm produced a sparse graph with 456 links (i.e., 9.3% of the original) with a −0.13 to 0.29 range for partial correlations. We were able to estimate the effective connectivity for 114 links (i.e., 25.0% of the total number of links in the sparse graph) using LOFS R3. The selected links and their directions were incorporated in the IC-network graph in Fig. 2 (for details on the functional circuits, see the Discussion).

3.10. IC modules: interaction

Fig. 3 shows the seven modules that were found, including their partial correlations. Of note, the posterior and anterior DMN were inversely correlated with the sensorimotor network, which is line with their original description as 'task-negative' networks (Fox et al., 2005). As regards its function, the visual-imagery/episodic-memory (VI/EM) module was strongly connected with the posterior DMN, thus seeming to combine efforts to integrate and uphold representations from brain-wide memory networks. The C-E-S module was most strongly connected with the anterior DMN and the sensorimotor network, showing very few connections with the other modules. For an anatomical overview of the modules found, a 3D rotating animation is available in the Supplementary materials (Figs. 6–12).

4. Discussion

This study explored the relationship between auditory verbal hallucinations (AVH) and the brain circuits involved in their mediation, using a model-free, network-based approach to analyze fMRI data obtained from 85 medication-resistant patients who were actively hallucinating. The analysis yielded 98 ICs of the brain, of which 18 correlated with the conscious experience of AVH. In addition, the 98 ICs clustered into seven modules with distinct and recognizable functions based on network metrics, a study of the literature, and a post-hoc association study. On the basis of these results, we created a network graph to provide a comprehensive overview of the brain's functions at the level of neural networks and to illustrate the networks' direct and indirect relationships with the mediation of AVH.

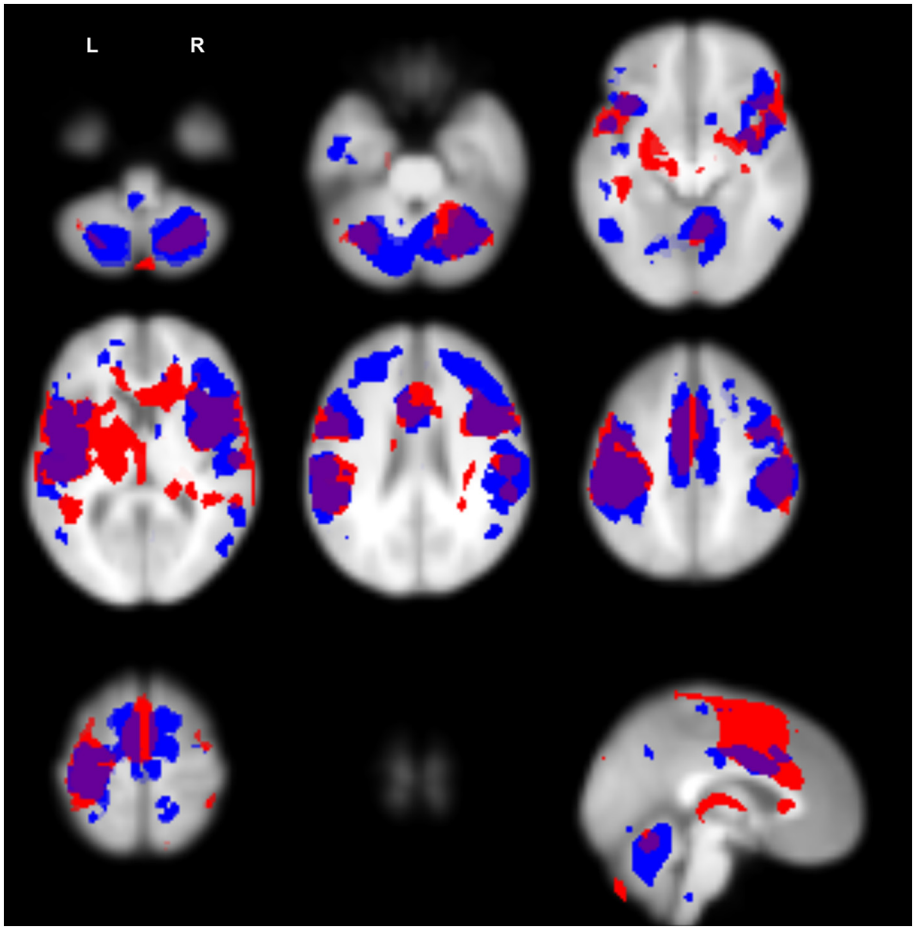

Fig. 4. Model-based vs model-free activation maps of AVH-related associated brain areas. Red color for the study by Looijestijn et al. (Looijestijn et al., 2013) using a symptom capture approach, blue color for the stacked ICs with significant positive beta's in the current study. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4.1. General architecture of the IC network

Using our model-free network approach, large-scale functional networks (such as the CEN, DMN, and SN), but also the cerebellum and other structures traditionally conceptualized as constituting single functional components, now appeared to fragment into smaller functional units. An explanation for the high level of detail of our ICA decomposition might be the high sampling rate of the PRESTO scan, which offers a fine delineation of functional (a)synchrony. Furthermore, a methodological study of FIX (Griffanti et al., 2014) has earlier reported on achieving higher dimensionality after the denoising-procedure. The high level of detail yielded by our ICA decomposition allowed comprehensive mapping of the brain's subfunctions involved in the mediation of AVH, whilst still acknowledging the brain as an extensively connected and intrinsically complex functional network (Smith et al., 2011). We made a systematic effort not to impose theoretical or pre-specified models onto our data. Therefore, it was noteworthy that the force-directed layout algorithm produced a relative positioning of ICs and IC modules that closely resemble the actual neuroanatomical positions of these areas in the adult human brain (Supplementary materials, Fig. 13). Overall, the high level of correspondence between the IC network and the anatomical network structure of the human brain, provided a first indication of the validity of our approach.

Major hubs were the left-sided and right-sided fronto-parietal networks (IC7, IC3), the precuneus and posterior cingulate (IC5), the thalamus and putamen (IC76, IC46), the hippocampus (IC93), the supramarginal gyrus (IC54), and a network surrounding the right temporo-occipital junction (IC58). This configuration is in strong accordance with the so-called 'rich club' of the human brain, a network of densely connected hubs thought to account for a large proportion (e.g., 80%) of information transfer within the brain (van den Heuvel & Sporns, 2011), thus providing a second validation of our approach.

4.2. Relationship between IC modules and auditory verbal hallucinations

The strongest links to AVH were found for ICs located within the sensorimotor (SM) module, the cognition, evaluation/salience, and response formation (C-E—S) module, and the cerebellar module (see below for further details per IC). Together they represent all the 11 ICs that significantly *activate* during AVH. In this study, DMN activity was found to be at a distance from hallucination-related regions, with the posterior DMN subdivisions mostly *deactivating* during AVH. Although functional hyperconnectivity and hyperactivity of DMN subdivisions are suggested to be essential processes for the occurrence of AVH (Whitfield-Gabrieli et al., 2011), our exploratory approach does not support that view. Bearing in mind the replicated findings concerning DMN hyperactivity in patients diagnosed with schizophrenia and their first-degree relatives (Whitfield-Gabrieli et al., 2009; Liu et al., 2010), this might be indicative of other types of psychopathology (i.e, not hallucinations). In their study, Jardri et al. (Jardri et al., 2013) found a comparable disengagement of the DMN during AVH and also found evidence for a role of spatial and temporal DMN instability in the emergence of AVH; therefore, their study is indicative of the complex constituents of AVH on multiple scale levels. In our study, apart from the anterior and posterior DMN regions, the brainstem and subcortical regions were mainly positioned at a distance from hallucination-related ICs and, therefore, appeared to have no significant role in the mediation of AVH. Also, the visual-imagery/episodic-memory module showed little or no relationship with AVH. This contradicts hypotheses suggesting that AVH have a source in unstable (episodic) memory (e.g., (para)hippocampal areas or putamen) (Waters et al., 2006; Amad et al.,

2013).

4.3. Relationship between individual ICs and auditory verbal hallucinations

The 11 ICs showing significantly positive relationships with hallucination timings comprise the cerebellum (both hemispheres and the vermis), the right anterior insula, Broca's homologue (right), the left pre- and postcentral gyri, the bilateral supramarginal gyrus, the medial frontal areas (including the anterior cingulate), the bilateral supplementary motor areas, and the bilateral frontal poles. These structures are often found in model-based studies of AVH (Looijestijn et al., 2013; Shergill et al., 2000; Sommer et al., 2008; Jardri et al., 2011; van Lutterveld et al., 2013), although several model-based studies also reported involvement of subcortical structures (Hoffman et al., 2011; Amad et al., 2013; Shergill et al., 2000). Fig. 4 shows the spatial maps of the *activated* ICs within a single brain, and contrasts these with the activation map from a previous study by our group based on model-based analyses of signal changes in AVH (Looijestijn et al., 2013). As shown in Fig. 4, these model-based and model-free activation maps largely overlap, with the model-based study yielding additional activity surrounding the thalamus and motor cortex in the activation map, and more extensive medial cerebellar, medial prefrontal, and fronto-polar activity in the stacked ICs from the present model-free study. The increased activation found by our previous model-based study in thalamus and motor cortex might be due to a number of factors. One explanation is the possibly superior power of the event-related approach to detect brain activity (due to the balloon presses) in these regions. In our model-free approach, the extensive activation of the medial cerebellum is of special interest in view of earlier studies that proposed a causal role for the cerebellum in psychosis and hallucinations (Andreasen & Pierson, 2008; Iglesia-Vaya et al., 2014; Powers et al., 2017). Powers et al. (Powers et al., 2017) found that the decreased activation of cerebellum corresponded with diminished belief-updating and rigidity of psychotic patients, and hypothesized that the cerebellum dysfunctions in updating top-down predictions; furthermore, they identified the superior temporal sulcus and the insula as discriminant regions that activate during the hallucinatory state. In a methodological study on the use of ICA to capture hallucinations, Leroy et al. (2017) (Leroy et al., 2017) report their hallucination-related ICs to be located in largely similar brain regions, e.g., cerebellar regions, insula, inferior parietal lobule, inferior frontal gyrus, middle frontal gyrus, cingulate gyrus, and inferior frontal medial and superior temporal regions. Notably, they also find multiple activation clusters in the medial cerebellum.

4.4. Functional circuits

The method we selected to estimate effective connectivity is among the most reliable available for detecting causal directions in fMRI data. Nevertheless, we could establish causal directions (arrows) for only 25% of our links. This posed limits to a more precise understanding of the functional circuits at hand. For instance, it prevented us from detecting a clear causal hierarchy between the various ICs, and from finding or excluding possible ‘Garden-of-Eden states’ (first movers) of hallucinations. Despite that limitation, Fig. 5 provides an overview of the ICs that emerged from our analysis. The circuit is built-up of AVH-activated ICs, and ICs that show a strong direct link with these ICs.

Of all ICs, left-sided motor cortex (IC1) showed the highest beta value, which most likely indicates the motor action coinciding with the balloon presses made by our study patients. The motor cortex was also found to have the strongest connection with somatosensory association cortices, including the auditory regions (IC19), which might represent the processing of language and/or the reciprocal feedback taking place during motor action (Amad et al., 2013). The bilateral superior parietal lobule (SPL, IC11) is strongly informed by the motor cortex (IC1), probably receiving input from the hand during balloon presses.

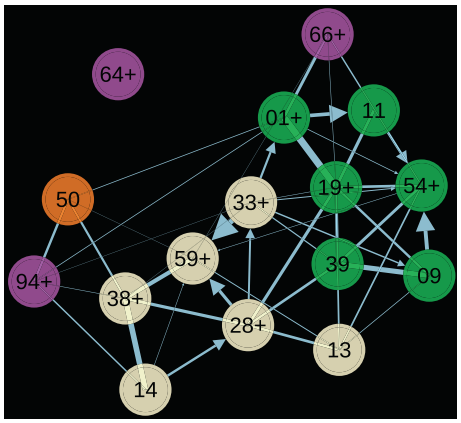

Fig. 5. Detail of IC network graph. ICs selected for significant positive beta's and identified tributaries. Partial correlations filtered at > 0.02. No direct links for IC64.

The SMA (IC33) - which provides input to motor cortex (IC1) - and the ACC (IC59) also showed strong activation during AVH. The SMA has been endowed with a number of functions, including motor preparation, the initiation of internally driven movement, and the processing of sequences of input within multiple domains (Alario et al., 2006; Cona & Semenza, 2017). The pre-SMA has been implicated in complex sequencing, ambiguity resolution, and task switching in relation to language (Hertrich et al., 2016). A study on the functional connectivity of the SMA shows pre-SMA output going to the posterior IFG, angular gyrus, and ACC (Kim et al., 2010), which matches with the effective connectivity found by us, albeit with a reversion of the reciprocal influence of pre-SMA and IFG. In a study directly hinting at a function of the SMA in mediating hallucinations, Clos et al. (Clos et al., 2013), observed increased connectivity between the left IFG, the SMA, and the insula in psychotic patients experiencing AVH, which they explained in terms of increased inner-speech generation.

The ACC has partially overlapping functions, such as cognitive control, salience, various top-down processes, and self-monitoring (Gasquoine, 2013; Ridderinkhof, 2004). All these functions may play a role in the mediation of AVH. In our analysis, we found that the ACC (IC59) receives input from the right-sided anterior insula and Broca's homologue (IC28), which matches with the study by Sridharan et al. (Sridharan et al., 2008), who found that a right-sided fronto-insular network strongly co-activated with the ACC (which is part of the SN). Together, these areas have an important causal role in modulating DMN and CEN activity which, as we saw, plays an important role in hybrid models of AVH mediation. The anterior insula has been proposed to function as the most crucial part of the SN, attributing ‘salience’ to intrinsic and extrinsic stimuli, and propelling this information forwards to the ACC for preferential attention in higher-order cognitive (prefrontal) areas interacting with the CEN or DMN (Menon & Uddin, 2010). Salience attribution seems to involve a process of risk assessment: the insula takes highly integrated (i.e., combined multimodal emotional and cognitive) information as its input and ‘calculates’ expected reward or punishment as a result of this input (Pezzulo et al., 2015). Predictions with maximum expected results are then prioritized within the ACC, and appropriate measures are prepared, e.g., whether to relax and enter into a resting/DMN state, or to become vigilant and enter an active/CEN state.

The right anterior insula and Broca's homologue (IC28) were also strongly connected with a prefrontal and SFG-focused component (IC38), although we were unable to establish any causal ties. Moreover, the SFG network includes the bilateral frontal pole, the temporal pole, the middle temporal gyrus, the cerebellum, and Broca's area. We hypothesize that this reflects working memory and associative functions following the conscious experience of AVH. The SFG network strongly

connects with the anterior medial SFG (IC10) and Broca's area (IC14). In our analysis, Broca's area was the strongest tributary, contributing to the activity levels of the combined Broca's homologue and right-sided insula (IC28). Interestingly, Sommer et al. (Sommer et al., 2008) reported on the relative inactivity of Broca's area, as compared to its homologue. Our findings are similar, but indicate that Broca's area does have a function in the mediation of AVH, albeit indirectly, e.g. by producing language that further 'upstream' is wrongly valued by the SN, and retained in working memory.

We found the right-sided posterior insula (IC39) to be connected with a range of AVH-related ICs, including the anterior insula. The posterior insula is thought to be functionally dissociated from the anterior insula, and to have a specific function in sensorimotor and somatosensory processing (Menon & Uddin, 2010; Chang et al., 2013). Correspondingly, we found the posterior insula to be strongly connected with auditory cortices (IC9). Therefore, we hypothesize that the links between the inferior parietal lobule (IC54) and the posterior insula (IC39) represent the back-projection of perceived hallucinations in the salience network, and the subsequent somatosensory registration of the ensuing percept.

In our analysis, the left-sided CEN (IC3) emerged as a hub that was inversely correlated with another hub, i.e., the right-sided IPL (IC54). Together with the inverse correlation with the SMA (IC33) of the left CEN (IC3), this suggests a relative *deactivation* of the lateralized CEN during AVH, although this was not found to be significant. The right-sided CEN (IC7) did deactivate significantly during AVH, acting as an interconnecting hub bridging the pDMN-module. The left-sided CEN (consistently found in resting-state studies) has been implicated in cognition and language processing, whereas the right-sided CEN is more often associated with somesthesis and action inhibition (Smith et al., 2009). This suggests that the abdication of lateralized cognitive control has a relation with the conscious experience of AVH. Rotarska-Jagiela et al. (Rotarska-Jagiela et al., 2010) found that the right-oriented CEN shows a disrupted intrinsic organization and a decreased rightward lateralization in patients experiencing AVH. Further away, we found that the pDMN hub (IC5) is inversely correlated with hallucination-activated ICs (IC19, IC28), but positively correlated with the lateralized CEN hubs. This partially conflicts with earlier studies that found inverse correlations between CEN and DMN, and the subsequent models that emphasize the competition between the CEN and DMN in their preponderance (Palaniyappan, 2012; Greicius & Menon, 2004; Chen et al., 2013). The deactivation of the DMN is proportional to the height of the cognitive demands of within-scanner tasks for subjects (Greicius & Menon, 2004). Therefore, in the present study, the correlation between the DMN and the lateralized CENs could indicate the relatively limited cognitive demands of the hallucination reporting. The bilateral CEN (IC13) correlated negatively with the pDMN hub (IC5), and was strongly connected with the hallucination-activated ICs (IC19, IC28, IC54, IC59). All this suggests that the CEN is functionally segregated into three subunits, with different relations to the aDMN, pDMN, SN, and hallucinatory activity. Therefore, further studies should take into account the functional separation of different CEN networks.

Lastly, a strong relation was found between the cerebellum and the patients' within-scanner balloon presses. Focusing on the cerebellar ICs that strongly link with other AVH-related ICs, the right cerebellum (IC94) was found to have a stronger connection with the ICs in the cognition, evaluation/salience and response formation (C-E-S) module, whereas the cerebellar vermis (IC66) had a stronger connection with the ICs of the sensorimotor module. Additionally, IC50 mainly constituting the cerebellum vermis was identified as a strong tributary in the hallucination circuit. The cerebellum is known to operate in a feed-forward system, e.g., in the computational processing of cerebral input, looping it back to the cerebrum with limited internal transmission. This function could be summarized as analyzing neural input for the prediction of a future sensory state, and the detection of discrepancies with actual signal patterns (i.e., sensory error prediction) (D'Angelo & Casali, 2012), thereby supporting timing and learning in mental processes. Hypothetically, when timing of cortical processes becomes desynchronized, this could amount to self-monitoring deficits of language networks and thus to AVH. However, it remains unclear whether the cerebellum is solely involved in the motor coordination of speech and the emotional modulation of speaking, or whether it also has a function in language (Ziegler, 2016).

4.5. Integration

Integrating the results of our analysis, we propose that the mediation and subsequent perception of AVH in the context of schizophrenia spectrum disorders relies on the involvement of medial prefrontal regions, the insula, the cerebellum, and the homologue of Broca's area. In this patient group, these components of the hallucination network appear to be essential in the mediation of AVH. The right-sided insula and Broca's homologue (IC28) are positioned centrally within the hallucination network, and appear to be responsible for the production of preconscious linguistic constructs to which excessive salience is assigned by the SN (IC28, IC59). The insula propagates the stimulus further downstream in the direction of middle frontal regions, where the SMA and ACC appear to respond to the false prediction, activating the planum temporale (which projects the voices into external space) (Looijestijn et al., 2013) and enforcing the conscious perception of AVH. Sustained attention to the ensuing percept is probably provided by the working memory ICs (IC38 and CENs). The bilateral CEN showed the strongest connection with these AVH-related ICs. The cerebellum, in turn, might be involved in the disrupted updating and learning from the false sensory predictions in psychosis or in a more physiological process, such as emotional modulation (Iglesia-Vaya et al., 2014; Powers et al., 2017). Two earlier studies focusing on the interaction between resting-state networks have underlined the importance of the SN in the mediation of AVH. Thus Manoliu et al. (2013) studied the triple network model, and reported that right anterior insular activity is associated with increased DMN-CEN interaction and severity of hallucinations, while Lefebvre et al. (2016) used network modelling to study effective connectivity in the triple network model plus the left hippocampus. Using an innovative design, they distinguished four phases in hallucinations, and found the SN (including the insula) to mediate switching between the DMN and CEN during active hallucinations. In addition, they found evidence that hippocampal input to the salience network evokes the hallucinatory state. In our results, DMN and CEN are not directly associated with the hallucinatory experience, rather showing deactivation during AVH. Altered influence between the DMN and CEN could still co-occur with or follow the onset of AVH within our patient group, however, this is hard to compare considering the breakdown of the large-scale DMN and CEN into multiple subnetworks in our analysis. The current study of RSNs in AVH similarly does not find evidence for direct involvement of the hippocampus in hallucination-onset. Instead of the hippocampus, our own analysis puts forth Broca's area and its homologue in the onset of AVH.

The mediation of AVH as outlined above aligns most strongly with established (model-based) hypotheses of AVH involving the disrupted self-monitoring of inner speech. Indeed, the central role of the anterior insula in falsely predicting threat or risk from harmless, internally generated narratives, fits the evolutionary model of psychosis as a partially adaptive strategy to increase the 'true positive' (successful) detection of threats (such as gossip or intrigue) at the cost of 'false positives' (hallucinations). Such a strategy may be rewarding in threatening situations that require high levels of vigilance or even a healthy amount of paranoia. In effect, the hypothesis presented here is a hybrid model that characterizes the mediation of AVH as a failure of source monitoring due to SN dysfunction, with subsequent processing of the internally mediated percepts as if they were mediated externally.

4.6. Limitations and suggestions for future research

The present study has several limitations. Firstly, the results of our model-free analysis depended on post-hoc statistical analyses and literature searches in order to attribute meaning to the reported ICs. Studies such as these need to consider the problem of 'reverse inference', i.e., (incorrectly) deducing the function of ICs on the basis of brain activation maps reported in previous studies (Poldrack, 2006). The criticism being that the attribution of a phenotypical function attributed to brain regions A in study X, on the basis of earlier research Y is invalid, as the brain regions A in study Y operate in a network of other brain regions and will probably have a role in multiple functions, in the sense that they are not specific to certain phenotypes. However, if the goal is to identify an optimal causal model for a phenotype within a broad context of neural mechanisms, this criticism is less relevant (Glymour & Hanson, 2016). Attributing psychological functions to specific brain regions is arbitrary and, instead, the focus should be on assessing the global mechanistic model of the symptom, disease or cognitive process under study. Secondly, a substantial number of ICs was found to originate primarily from noise, and had to be removed using denoising algorithms based on supervised learning. The substantial amount of noise in the data might have been due to the liberal method of temporal filtering used or, alternatively, to the possibly lower signal-to-noise ratio of the PRESTO scanning technique (Hesselmann et al., 2004). Nevertheless, the higher sampling rate of PRESTO (in comparison with Echo Planar Imaging) was a clear advantage of our study with regard to power and the ability to discern potentially meaningful signals from physiological noise (van Gelderen et al., 2012). In the third place, the absence of a control group prevented us from making any firm inferences regarding the specific pathological deficits or pathogenetic mechanisms involved in the mediation of hallucinations. However, it should be emphasized that the primary goal of this study was not to investigate such pathogenetic mechanisms, but to identify the mechanistic model for the occurrence of AVH within the context of the afflicted brain: more specifically, to facilitate local intervention with the aid of repetitive transcranial magnetic stimulation or transcranial direct current stimulation. Pathological changes in psychosis may significantly alter normal functional neuroanatomy (both in terms of structure and function) (Curcic-Blake et al., 2017; Shepherd et al., 2012). Therefore, for the purpose of clinical intervention, it is necessary to know the neural mechanisms underlying psychotic experiences within the context of the disease, rather than the healthy situation. Additionally, our study shows that many functionally significant ICs are not significantly associated with AVH experience, providing a within-subject control condition. Future studies that aim to examine pathogenic mechanisms might benefit from including a control group consisting of healthy controls, siblings of psychotic patients, non-hallucinating patients with previous episodes of psychosis, or patients with infrequent hallucinations. Such imaging studies can be performed during the resting state or during an auditory stimulus-detection task that is matched to the psychotic experiences of the patients. In the fourth place, as noted above, the LOFS R3 algorithm used allowed to determine causal directions for only 25% of the links. As LOFS R3 does not make any forced choices, this may indicate the existence of reciprocal connections in the remaining links, as can be expected in neural networks (Mumford & Ramsey, 2014). On the other hand, it may also indicate that this algorithm prefers accuracy of directionality over directionality. One way to improve this estimation would be to i) choose components of interest based (COIs) on the present study, ii) use the time series of COIs in relation to the modeled BOLD response according to the hallucination timings to estimate the order of activation of COIs, and iii) pre-inform the causal search algorithm with directions of links based on this order of activation. Although this would only be feasible in individual subject-level analyses of effective connectivity, the benefit would be substantial as a detailed depiction of functional circuits involved in the mediation of AVH would allow accurate predictions of locations to be targeted with therapeutic techniques within individuals, i.e., by selecting the most influential nodes in the functional networks at hand (Jia et al., 2013). Moreover, enforcing further sparsity in the network might also help to estimate causal links, as this will reduce cyclical (feedback) connections and benefit the display of the dominant direction of information in the brain circuits under study.

4.7. Conclusions

We systematically decomposed the fMRI data from the hallucinating brains of patients diagnosed with a schizophrenia spectrum disorder into functional subnetworks, and reconstructed these into a whole-brain directed network. This method, which we compared to the draining of a pond to lay bare its entire ecosystem, revealed 98 independent components (ICs) which were active in patients who had consciously experienced AVH during the time of scanning. These ICs clustered into seven modules with distinct physiological functions, involving resting state, central-executive, salience, cerebellar, subcortical, and stimulus-response processing. Functional subnetworks comprising the hallucination network are Broca's right homologue, the right insula, the bilateral anterior cingulate, premotor cortex, and the supramarginal gyrus, whereas the CENs, Broca's area, and cerebellar regions constitute probable and more distant tributaries to the mediation of AVH. On the basis of the present findings, we conclude that in our treatment-refractory patient group, AVH appear to be largely mediated by the SN making false predictions about the risk and (hence) origin of linguistic percepts derived from Broca's homologue, followed by subsequent processing errors in the anterior cingulate gyrus, cerebellum and other cognitive areas. Our findings mostly comply with model-based studies reporting faulty error monitoring as a major factor for the mediation of AVH. Future local intervention studies should consider focusing interventions on Broca's homologue or on SN subparts, anterior insula, and anterior cingulate cortex, instead of the traditional left temporoparietal cortex (T3P3 in EEG electrode placement system) (Slotema et al., 2010).

Supplementary data to this article can be found online at https://doi.org/10.1016/j.nicl.2018.09.016.

Acknowledgements

This study was funded by the Parnassia Psychiatric Institute.

Declarations of interest

None.

References

Alario, F.X., Chainay, H., Lehericy, S., Cohen, L., 2006. The role of the supplementary motor area (SMA) in word production. Brain Res. 1076 (1), 129–143. https://doi.org/10.1016/j.brainres.2005.11.104.

Amad, A., Cachia, A., Gorwood, P., Pins, D., Delmaire, C., Rolland, B., et al., 2013. The multimodal connectivity of the hippocampal complex in auditory and visual hallucinations. Mol. Psychiatry 19 (2), 1–8. https://doi.org/10.1038/mp.2012.181.

Anderson, T.W., Darling, D.A., 1952. Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. Ann. Math. Stat. https://doi.org/10.2307/2236446.

Andreasen, N.C., Pierson, R., 2008. The Role of the Cerebellum in Schizophrenia. Biol. Psychiatry 64 (2), 81–88. https://doi.org/10.1016/j.biopsych.2008.01.003.

Andreasen, N.C., Flaum, M., Arndt, S., 1992. The comprehensive Assessment of Symptoms and history (CASH): an instrument for assessing diagnosis and psychopathology. Arch. Gen. Psychiatry 49 (8), 615–623. https://doi.org/10.1001/archpsyc.1992.01820080023004.

Andrews-Hanna, J.R., Reidler, J.S., Sepulcre, J., Poulin, R., Buckner, R.L., 2010. Functional-anatomic fractionation of the brain's default network. Neuron 65 (4), 550–562. https://doi.org/10.1016/j.neuron.2010.02.005.

Apps, R., Garwicz, M., 2005. Anatomical and physiological foundations of cerebellar information processing. Nat. Rev. Neurosci. 6 (4), 297–311. https://doi.org/10.1038/nrn1646.

Bechara, A., Damasio, H., Damasio, A.R., 2000. Emotion, decision making and the

orbitofrontal cortex. Cereb. Cortex 10 (3), 295–307.
Beckmann, C.F., Smith, S.M., 2004. Probabilistic Independent Component Analysis for Functional magnetic Resonance Imaging. IEEE Trans. Med. Imaging 23 (2), 137–152. https://doi.org/10.1109/TMI.2003.822821.
Blom, J.D., 2015. Auditory hallucinations. Handb. Clin. Neurol. 129, 433–455. https://doi.org/10.1016/B978-0-444-62630-1.00024-X.
Blondel, V.D., 2008. Fast unfolding of communities in large networks. J. Stat. Mech. 1–12.
Chang, L.J., Yarkoni, T., Khaw, M.W., Sanfey, A.G., 2013. Decoding the role of the insula in human cognition: functional parcellation and large-scale reverse inference. Cereb. Cortex 23 (3), 739–749. https://doi.org/10.1093/cercor/bhs065.
Chen, J., Chen, Z., 2012. Extended BIC for small-n-large-P sparse GLM. Stat. Sin. 22 (2), 1–28. https://doi.org/10.5705/ss.2010.216.
Chen, A.C., Oathes, D.J., Chang, C., Bradley, T., Zhou, Z.-W., Williams, L.M., et al., 2013. Causal interactions between fronto-parietal central executive and default-mode networks in humans. Proc. Natl. Acad. Sci. U. S. A. 110 (49), 19944–19949.
Clos, M., Diederen, K.M.J., Meijering, A.L., Sommer, I.E., Eickhoff, S.B., 2013. Aberrant connectivity of areas for decoding degraded speech in patients with auditory verbal hallucinations. Brain Struct. Funct. 219 (2), 581–594. https://doi.org/10.1007/s00429-013-0519-5.
Cona, G., Semenza, C., 2017. Supplementary motor area as key structure for domain-general sequence processing: a unified account. Neurosci. Biobehav. Rev. 72, 28–42. https://doi.org/10.1016/j.neubiorev.2016.10.033.
Curcic-Blake, B., Ford, J.M., Hubl, D., Orlov, N.D., Sommer, I.E., Waters, F., et al., 2017. Interaction of language, auditory and memory brain networks in auditory verbal hallucinations. Prog. Neurobiol. 148, 1–20. https://doi.org/10.1016/j.pneurobio.2016.11.002.
D'Angelo, E., Casali, S., 2012. Seeking a unified framework for cerebellar function and dysfunction: from circuit operations to cognition. Front. Neural Circ. 6, 116. https://doi.org/10.3389/fncir.2012.00116.
David, O., Guillemain, I., Saillet, S., Reyt, S., Deransart, C., Segebarth, C., Depaulis, A., 2008. Identifying neural drivers with functional MRI: an electrophysiological validation. PLoS Biol. 6 (12), 2683–2697. https://doi.org/10.1371/journal.pbio.0060315.
Diederen, K.M.J., Neggers, S.F.W., Daalman, K., Blom, J.D., Goekoop, R., Kahn, R.S., Sommer, I.E.C., 2010. Deactivation of the parahippocampal gyrus preceding auditory hallucinations in Schizophrenia. Am. J. Psychiatr. 167 (4), 427–435. https://doi.org/10.1176/appi.ajp.2009.09040456.
Dobryakova, E., Boukrina, O., Wylie, G.R., 2015. Investigation of information flow during a novel working memory task in individuals with traumatic brain injury. Brain Connect. 5 (7), 433–441. https://doi.org/10.1089/brain.2014.0283.
Epskamp, S., Cramer, A., Waldorp, L.J., 2012. qgraph: Network visualizations of relationships in psychometric data. J. Stat. Softw. 48 (4). https://doi.org/10.18637/jss.v048.i04.
Filippini, N., MacIntosh, B.J., Hough, M.G., Goodwin, G.M., Frisoni, G.B., Smith, S.M., et al., 2009. Distinct patterns of brain activity in young carriers of the APOE-ε4 allele. Proc. Natl. Acad. Sci. 106 (17), 7209–7214. https://doi.org/10.1073/pnas.0811879106.
Fox, M.D., Snyder, A.Z., Vincent, J.L., Corbetta, M., Van Essen, D.C., Raichle, M.E., 2005. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proc. Natl. Acad. Sci. 102 (27), 9673–9678. https://doi.org/10.1073/pnas.0504136102.
Foygel, R., Drton, M., 2010. Extended Bayesian information criteria for Gaussian graphical models. In: Lafferty, J.D., Williams, C.K.I., Shawe-Taylor, J., Zemel, R.S., Culotta, A. (Eds.), Advances in Neural Information Processing Systems. Vol. 23. Curran Associates, Inc., pp. 604–612.
Friedman, J., Hastie, T., Tibshirani, R., 2008. Sparse inverse covariance estimation with the graphical lasso. Biostatistics (Oxford, England) 9 (3), 432–441. https://doi.org/10.1093/biostatistics/kxm045.
Friston, K.J., Harrison, L., Penny, W., 2003. Dynamic causal modelling. NeuroImage 19 (4), 1273–1302. https://doi.org/10.1016/S1053-8119(03)00202-7.
Gasquoine, P.G., 2013. Localization of function in anterior cingulate cortex: from psychosurgery to functional neuroimaging. Neurosci. Biobehav. Rev. 37 (3), 340–348. https://doi.org/10.1016/j.neubiorev.2013.01.002.
Glymour, C., Hanson, C., 2016. Reverse inference in neuropsychology. Br. J. Philos. Sci. 67 (4), 1139–1153. https://doi.org/10.1093/bjps/axv019.
Greicius, M.D., Menon, V., 2004. Default-mode activity during a passive sensory task: uncoupled from deactivation but impacting activation. J. Cogn. Neurosci. 16 (9), 1484–1492. https://doi.org/10.1162/0898929042568532.
Griffanti, L., Salimi-Khorshidi, G., Beckmann, C.F., Auerbach, E.J., Douaud, G., Sexton, C.E., et al., 2014. ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. NeuroImage 95, 232–247. https://doi.org/10.1016/j.neuroimage.2014.03.034.
Hertrich, I., Dietrich, S., Ackermann, H., 2016. The role of the supplementary motor area for speech and language processing. Neurosci. Biobehav. Rev. 68, 602–610. https://doi.org/10.1016/j.neubiorev.2016.06.030.
Hesselmann, V., Girnus, R., Wedekind, C., Hunsche, S., Bunke, J., Schulte, O., et al., 2004. Functional MRI using multiple receiver coils: BOLD signal changes and signal-to-noise ratio for three-dimensional-PRESTO vs. single shot EPI in comparison to a standard quadrature head coil. J. Magn. Reson. Imaging 20 (2), 321–326. https://doi.org/10.1002/jmri.20101.
Hoffman, R.E., Pittman, B., Constable, R.T., Bhagwagar, Z., Hampson, M., 2011. Time course of regional brain activity accompanying auditory verbal hallucinations in schizophrenia. Br. J. Psychiatry 198 (4), 277–283. https://doi.org/10.1192/bjp.bp.110.086835.
Iglesia-Vaya, la, De, M., Escartí, M.J., Molina-Mateo, J., Martí-Bonmatí, L., Gadea, M., Castellanos, F.X., et al., 2014. Abnormal synchrony and effective connectivity in patients with schizophrenia and auditory hallucinations. NeuroImage. Clin. 6, 171–179. https://doi.org/10.1016/j.nicl.2014.08.027.
Jardri, R., Pouchet, A., Pins, D., Thomas, P., 2011. Cortical activations during auditory verbal hallucinations in schizophrenia: a coordinate-based meta-analysis. Am. J. Psychiatry 168 (1), 73–81. https://doi.org/10.1176/appi.ajp.2010.09101522.
Jardri, R., Thomas, P., Delmaire, C., Delion, P., Pins, D., 2013. The neurodynamic organization of modality-dependent hallucinations. Cereb. Cortex 23 (5), 1108–1117. https://doi.org/10.1093/cercor/bhs082.
Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. NeuroImage 17 (2), 825–841. https://doi.org/10.1006/nimg.2002.1132.
Jia, T., Liu, Y.-Y., Csóka, E., Pósfai, M., Slotine, J.-J., Barabási, A.-L., 2013. Emergence of bimodality in controlling complex networks. Nat. Commun. 4, 2002. https://doi.org/10.1038/ncomms3002.
Kim, J.-H., Lee, J.-M., Jo, H.J., Kim, S.H., Lee, J.H., Kim, S.T., et al., 2010. Defining functional SMA and pre-SMA subregions in human MFC using resting state fMRI: Functional connectivity-based parcellation method. NeuroImage 49 (3), 2375–2386. https://doi.org/10.1016/j.neuroimage.2009.10.016.
Lefebvre, S., Demeulemeester, M., Leroy, A., Delmaire, C., Lopes, R., Pins, D., et al., 2016. Network dynamics during the different stages of hallucinations in schizophrenia. Hum. Brain Mapp. 37 (7), 2571–2586. https://doi.org/10.1002/hbm.23197.
Leroy, A., Foucher, J.R., Pins, D., Delmaire, C., Thomas, P., Roser, M.M., et al., 2017. fMRI capture of auditory hallucinations: validation of the two-steps method. Hum. Brain Mapp. 38 (10), 4966–4979. https://doi.org/10.1002/hbm.23707.
Liu, H., Kaneko, Y., Ouyang, X., Li, L., Hao, Y., Chen, E.Y.H., et al., 2010. Schizophrenic patients and their unaffected siblings share increased resting-state connectivity in the task-negative network but not its anticorrelated task-positive network. Schizophr. Bull. 38 (2), 285–294. https://doi.org/10.1093/schbul/sbq074.
Liu, Y.-Y., Slotine, J.-J., Barabási, A.-Á., 2011. Controllability of complex networks. Nature 473 (7346), 167–173. https://doi.org/10.1038/nature10011.
Loh, M., Rolls, E., Deco, G., 2007. Statistical fluctuations in attractor networks related to Schizophrenia. Pharmacopsychiatry 40 (S 1), S78–S84. https://doi.org/10.1055/s-2007-990304.
Looijestijn, J., Diederen, K.M.J., Goekoop, R., Sommer, I.E.C., Daalman, K., Kahn, R.S., et al., 2013. The auditory dorsal stream plays a crucial role in projecting hallucinated voices into external space. Schizophr. Res. 146 (1–3), 1–6. https://doi.org/10.1016/j.schres.2013.02.004.
Looijestijn, J., Blom, J.D., Aleman, A., Hoek, H.W., Goekoop, R., 2015. An integrated network model of psychotic symptoms. Neurosci. Biobehav. Rev. 59 (C), 238–250. https://doi.org/10.1016/j.neubiorev.2015.09.016.
Manelis, A., Almeida, J.R.C., Stiffler, R., Lockovich, J.C., Aslam, H.A., Phillips, M.L., 2016. Anticipation-related brain connectivity in bipolar and unipolar depression: a graph theory approach. Brain 139 (Pt 9), 2554–2566. https://doi.org/10.1093/brain/aww157.
Manoliu, A., Riedl, V., Zherdin, A., Mühlau, M., Schwerthöffer, D., Scherr, M., et al., 2013. Aberrant dependence of default mode/central executive network interactions on anterior insular salience network activity in Schizophrenia. Schizophr. Bull. 40 (2), 428–437. https://doi.org/10.1093/schbul/sbt037.
Menon, V., Uddin, L.Q., 2010. Saliency, switching, attention and control: a network model of insula function. Brain Struct. Funct. 214 (5–6), 655–667. https://doi.org/10.1007/s00429-010-0262-0.
Mumford, J.A., Ramsey, J.D., 2014. Bayesian networks for fMRI: a primer. NeuroImage 86 (C), 573–582. https://doi.org/10.1016/j.neuroimage.2013.10.020.
Neggers, S.F.W., Hermans, E.J., Ramsey, N.F., 2008. Enhanced sensitivity with fast three-dimensional blood-oxygen-level-dependent functional MRI: comparison of SENSE-PRESTO and 2D-EPI at 3 T. NMR Biomed. 21 (7), 663–676. https://doi.org/10.1002/nbm.1235.
Nieuwenhuis, I.L.C., Takashima, A., 2011. The role of the ventromedial prefrontal cortex in memory consolidation. Behav. Brain Res. 218 (2), 325–334. https://doi.org/10.1016/j.bbr.2010.12.009.
Northoff, G., Qin, P., 2010. How can the brain's resting state activity generate hallucinations? A 'resting state hypothesis' of auditory verbal hallucinations. Schizophr. Res. 127 (1–3), 1–13. https://doi.org/10.1016/j.schres.2010.11.009.
Ohayon, M.M., 2000. Prevalence of hallucinations and their pathological associations in the general population. Psychiatry Res. 97 (2–3), 153–164. https://doi.org/10.1016/S0165-1781(00)00227-4.
Opsahl, T., Agneessens, F., Skvoretz, J., 2010. Node centrality in weighted networks: generalizing degree and shortest paths. Soc. Networks 32 (3), 245–251. https://doi.org/10.1016/j.socnet.2010.03.006.
Palaniyappan, L., 2012. Does the salience network play a cardinal role in psychosis? An emerging hypothesis of insular dysfunction. J. Psychiatry Neurosci. 37 (1), 17–27. https://doi.org/10.1503/jpn.100176.
Pezzulo, G., Barca, L., Friston, K.J., 2015. Active inference and cognitive-emotional interactions in the brain. Behav. Brain Sci. 38, e85. https://doi.org/10.1017/S0140525X14001009.
Poldrack, R., 2006. Can cognitive processes be inferred from neuroimaging data? Trends Cogn. Sci. 10 (2), 59–63. https://doi.org/10.1016/j.tics.2005.12.004.
Powers, A.R., Mathys, C., Corlett, P.R., 2017. Pavlovian conditioning–induced hallucinations result from overweighting of perceptual priors. Science 357 (6351), 596–600. https://doi.org/10.1126/science.aan3458.
Pruessmann, K.P., Weiger, M., Scheidegger, M.B., Boesiger, P., 1999. SENSE: sensitivity encoding for fast MRI. Magn. Reson. Med. 42 (5), 952–962.
Ramsey, J.D., Hanson, S.J., Hanson, C., Halchenko, Y.O., Poldrack, R.A., Glymour, C., 2010. Six problems for causal inference from fMRI. NeuroImage 49 (2), 1545–1558. https://doi.org/10.1016/j.neuroimage.2009.08.065.
Ramsey, J.D., Hanson, S.J., Glymour, C., 2011. Multi-subject search correctly identifies

causal connections and most causal directions in the DCM models of the Smith et al. simulation study. NeuroImage 58 (3), 838–848. https://doi.org/10.1016/j.neuroimage.2011.06.068.
Ramsey, J.D., Sanchez-Romero, R., Glymour, C., 2014. Non-Gaussian methods and high-pass filters in the estimation of effective connections. NeuroImage 84 (C), 986–1006. https://doi.org/10.1016/j.neuroimage.2013.09.062.
Ridderinkhof, K.R., 2004. The role of the medial frontal cortex in cognitive control. Science 306 (5695), 443–447. https://doi.org/10.1126/science.1100301.
Rotarska-Jagiela, A., van de Ven, V., Oertel-Knöchel, V., Uhlhaas, P.J., Vogeley, K., Linden, D.E.J., 2010. Resting-state functional network correlates of psychotic symptoms in schizophrenia. Schizophr. Res. 117 (1), 21–30. https://doi.org/10.1016/j.schres.2010.01.001.
Rubinov, M., Sporns, O., 2011. Weight-conserving characterization of complex functional brain networks. NeuroImage 56 (4), 2068–2079. https://doi.org/10.1016/j.neuroimage.2011.03.069.
Salimi-Khorshidi, G., Douaud, G., Beckmann, C.F., Glasser, M.F., Griffanti, L., Smith, S.M., 2014. Automatic denoising of functional MRI data: Combining independent component analysis and hierarchical fusion of classifiers. NeuroImage 90, 449–468. https://doi.org/10.1016/j.neuroimage.2013.11.046.
Schmahmann, J.D., Sherman, J.C., 1998. The cerebellar cognitive affective syndrome. Brain 121, 561–579 Pt 4.
Seth, A.K., Barrett, A.B., Barnett, L., 2015. Granger causality analysis in neuroscience and neuroimaging. J. Neurosci. 35 (8), 3293–3297. https://doi.org/10.1523/JNEUROSCI.4399-14.2015.
Shepherd, A.M., Laurens, K.R., Matheson, S.L., Carr, V.J., Green, M.J., 2012. Systematic meta-review and quality assessment of the structural brain alterations in schizophrenia. Neurosci. Biobehav. Rev. 36 (4), 1342–1356 (Elsevier Ltd).
Shergill, S.S., 2004. Temporal course of auditory hallucinations. Br. J. Psychiatry 185 (6), 516–517. https://doi.org/10.1192/bjp.185.6.516.
Shergill, S.S., Brammer, M.J., Williams, S.C.R., Murray, R.M., McGuire, P.K., 2000. Mapping auditory hallucinations in schizophrenia using functional magnetic resonance imaging. Arch. Gen. Psychiatry 57 (11), 1033–1038. https://doi.org/10.1001/archpsyc.57.11.1033.
Slotema, C.W., Blom, J.D., Hoek, H.W., Sommer, I.E.C., 2010. Should we expand the toolbox of psychiatric treatment methods to include Repetitive Transcranial Magnetic Stimulation (rTMS)? A meta-analysis of the efficacy of rTMS in psychiatric disorders. J. Clin. Psychiatry 71 (7), 873–884. https://doi.org/10.4088/JCP.08m04872gre.
Smith, S.M., Fox, P.T., Miller, K.L., Glahn, D.C., Fox, P.M., MacKay, C.E., et al., 2009. Correspondence of the brain's functional architecture during activation and rest. Proc. Natl. Acad. Sci. 106 (31), 13040–13045. https://doi.org/10.1073/pnas.0905267106.
Smith, S.M., Miller, K.L., Salimi-Khorshidi, G., Webster, M., Beckmann, C.F., Nichols, T.E., et al., 2011. Network modelling methods for FMRI. NeuroImage 54 (2), 875–891. https://doi.org/10.1016/j.neuroimage.2010.08.063.
Sommer, I.E.C., Diederen, K.M.J., Blom, J.D., Willems, A., Kushan, L., Slotema, K., et al., 2008. Auditory verbal hallucinations predominantly activate the right inferior frontal area. Brain 131 (12), 3169–3177. https://doi.org/10.1093/brain/awn251.
Sridharan, D., Levitin, D.J., Menon, V., 2008. A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. Proc. Natl. Acad. Sci. 105 (34), 12569–12574. https://doi.org/10.1073/pnas.0800005105.
van Borkulo, C.D., Borsboom, D., Epskamp, S., Blanken, T.F., Boschloo, L., Schoevers, R.A., Waldorp, L.J., 2014. A new method for constructing networks from binary data. Sci. Rep. 4 (1), 10–14. https://doi.org/10.1038/srep05918.
van de Ven, V.G., Formisano, E., Röder, C.H., Prvulovic, D., Bittner, R.A., Dietz, M.G., et al., 2005. The spatiotemporal pattern of auditory cortical responses during verbal hallucinations. NeuroImage 27 (3), 644–655. https://doi.org/10.1016/j.neuroimage.2005.04.041.
van den Heuvel, M.P., Sporns, O., 2011. Rich-club organization of the human connectome. J. Neurosci. 31 (44), 15775–15786. https://doi.org/10.1523/JNEUROSCI.3539-11.2011.
Van Dijk, K.R.A., Sabuncu, M.R., Buckner, R.L., 2012. The influence of head motion on intrinsic functional connectivity MRI. NeuroImage 59 (1), 431–438. https://doi.org/10.1016/j.neuroimage.2011.07.044.
van Gelderen, P., Duyn, J.H., Ramsey, N.F., Liu, G., Moonen, C.T.W., 2012. The PRESTO technique for fMRI. NeuroImage 62 (2), 676–681. https://doi.org/10.1016/j.neuroimage.2012.01.017.
van Lutterveld, R., Diederen, K.M.J., Koops, S., Begemann, M.J.H., Sommer, I.E.C., 2013. The influence of stimulus detection on activation patterns during auditory hallucinations. Schizophr. Res. 145 (1–3), 27–32. https://doi.org/10.1016/j.schres.2013.01.004.
Waters, F., Badcock, J., Michie, P., Maybery, M., 2006. Auditory hallucinations in schizophrenia: intrusive thoughts and forgotten memories. Cogn. Neuropsychiatry 11 (1), 65–83. https://doi.org/10.1080/13546800444000191.
Whitfield-Gabrieli, S., Thermenos, H.W., Milanovic, S., Tsuang, M.T., Faraone, S.V., McCarley, R.W., et al., 2009. Hyperactivity and hyperconnectivity of the default network in schizophrenia and in first-degree relatives of persons with schizophrenia. Proc. Natl. Acad. Sci. 106 (4), 1279–1284. https://doi.org/10.1073/pnas.0809141106.
Whitfield-Gabrieli, S., Moran, J.M., Nieto-Castañón, A., Triantafyllou, C., Saxe, R., Gabrieli, J.D.E., 2011. Associations and dissociations between default and self-reference networks in the human brain. NeuroImage 55 (1), 225–232. https://doi.org/10.1016/j.neuroimage.2010.11.048.
Woods, S.W., 2003. Chlorpromazine equivalent doses for the newer atypical antipsychotics. J. Clin. Psychiatry 64 (6), 663–667.
Ziegler, W. (2016). The phonetic cerebellum. In The Linguistic Cerebellum (pp. 1–32). Elsevier. https://doi.org/10.1016/B978-0-12-801608-4.00001-3.