Schizophrenia Bulletin vol. 49 no. suppl. 1 pp. S58–S67, 2023
https://doi.org/10.1093/schbul/sbac028
Advance Access publication May 21, 2022

SUPPLEMENT ARTICLE

Neural Activation in the Ventromedial Prefrontal Cortex Precedes Conscious Experience of Being in or out of a Transient Hallucinatory State

Kenneth Hugdahl[1,2,3,*], Alexander R. Craven[1,4], Erik Johnsen[2,5,6], Lars Ersland[1,4], Drozdstoy Stoyanov[7], Sevdalina Kandilarova[7], Lydia Brunvoll Sandøy[1], Rune A. Kroken[2,5], Else-Marie Løberg[2,5,8,9], and Iris E. C. Sommer[10]

[1]Department of Biological and Medical Psychology, University of Bergen, Bergen, Norway; [2]Division of Psychiatry, Haukeland University Hospital, Bergen, Norway; [3]Department of Radiology, Haukeland University Hospital, Bergen, Norway; [4]Department of Clinical Engineering, Haukeland University Hospital, Bergen, Norway; [5]NORMENT Center for the Study of Mental Disorders, Haukeland University Hospital, Bergen, Norway; [6]Department of Clinical Medicine, University of Bergen, Bergen, Norway; [7]Department of Psychiatry and Medical Psychology, and Research Institute, Medical University of Plovdiv, Plovdiv, Bulgaria; [8]Department of Addiction Medicine, Haukeland University Hospital, Bergen, Norway; [9]Department of Clinical Psychology, University of Bergen, Bergen, Norway; [10]Rijks Universiteit Groningen (RUG), Department of Biomedical Sciences of Cells and Systems and Department of Psychiatry, University Medical Center Groningen (UMCG), Netherlands

*To whom correspondence should be addressed; IBMP, University of Bergen, Jonas Lies vei 91, 5009 Bergen, Norway; tel: +47-91181062, e-mail: hugdahl@uib.no

Background and Hypotheses: **Auditory verbal hallucinations (AVHs) is not only a common symptom in schizophrenia but also observed in individuals in the general population. Despite extensive research, AVHs are poorly understood, especially their underlying neuronal architecture. Neuroimaging methods have been used to identify brain areas and networks that are activated *during* hallucinations. A characteristic feature of AVHs is, however, that they fluctuate over time, with varying frequencies of starts and stops. An unanswered question is, therefore, what neuronal events co-occur with the initiation and inhibition of an AVH episode. *Study Design:* We investigated brain activation with fMRI in 66 individuals who experienced multiple AVH-episodes while in the scanner. We extracted time-series fMRI-data and monitored changes second-by-second from 10 s before to 15 s after participants indicated the start and stop of an episode, respectively, by pressing a hand-held response-button. *Study Results:* We found a region in the ventromedial prefrontal cortex (VMPFC) which showed a significant increase in activation initiated a few seconds before participants indicated the start of an episode, and a corresponding decrease in activation initiated a few seconds before the end of an episode. *Conclusions:* The consistent increase and decrease in activation in this area in advance of the consciously experienced presence or absence of the "voice" imply that this region may act as a switch in turning episodes on and off. The activation is unlikely to be confounded by motor responses. The findings could have clinical implications for brain stimulation treatments, like transcranial magnetic stimulation.**

Key words: fMRI/Schizophrenia/Auditory verbal hallucinations/Non-clinical hallucinations/Ventromedial prefrontal cortex/Button-press

Introduction

Auditory verbal hallucinations (AVHs) or "hearing voices" in the absence of a corresponding auditory source, is a remarkable state of the mind. AVHs are traditionally seen as a typical symptom of schizophrenia,[1–7] but also commonly occur in other psychiatric and neurological disorders.[8–10] We know that AVHs have neuronal correlates, as observed in hemodynamic (fMRI and PET) and electrophysiology (EEG and MEG) studies (for meta-analyses,[11–13] and for reviews[14–18]). However, these studies have mostly been focused on patterns of neuronal activation *during* a hallucinatory episode and have thus not addressed the question of why AVHs fluctuate over time. Thus, although hallucinatory episodes show great variation across individuals with regard to

frequency and duration of episodes, they are typically not continuously present. Why is that? To advance our knowledge of the underlying neuronal mechanisms of the subjectively perceived fluctuating characteristic of AVHs, it would be crucial to understand how the brain transits between hallucinatory and non-hallucinatory states. One could think of the transition as a balance between excitatory and inhibitory forces[6,19–21] with excess of excitatory forces resulting in an imbalance that triggers the onset of an episode, which occasionally comes back into balance, resulting in the offset of an episode. Tracking the temporal characteristics of AVHs may therefore lead to a breakthrough not only in our theoretical understanding of this remarkable mental phenomenon, but also which could open new leads for clinical therapeutic approaches. Brain activation can be monitored on-line with functional magnetic resonance imaging (fMRI), where changes in neuronal activation are estimated from modeling of the blood-oxygenation-dependent (BOLD) function.[22] Although being an indirect measure of neuronal activity, fMRI has the advantage of spatial specificity compared to other methods. Monitoring AVH fluctuations with fMRI is a very demanding paradigm, as it requires participants to be aware and thoughtful of their AVHs, to lay still in the scanner while experiencing AVHs and indicating AVH-episode onsets and offsets with a button-press.[11,13,23–25] In addition, informative functional scans can only be obtained when participants experience a required minimum of episodes during the scanning session, as neither continuous hallucinations during the session, nor a session with too few hallucinations will allow for the study of the neuronal time-course of episode on- and offsets. Also, some luck is necessary as about a total of 50 functional MRI scans are needed to "catch" spontaneous and unpredictable hallucinatory episodes. Only a few research previous studies exist, and they have typically been based on less than 10 participants.[26–28] However, to study on- and offsets, larger numbers of subjects and scans are critical, as a hypothesized switch in activation between presence or absence of an episode will occur in blocks of a few seconds at most. To increase power to detect such subtle changes in brain activation during brief moments of on and offsets, we joined forces from three research groups in three different countries to recruit, in this context, a sufficiently large sample of participants hallucinating while in the scanner. As hallucinatory experiences cross the border between pathological and normal conditions, we included both clinical and nonclinical "voice-hearers" in the sample, focusing on tracking the neuronal signatures of AVH-fluctuations per se, not restricted to a particular diagnostic group or mental condition. Before analyzing the time-course around on- and off-events, we first had to confirm that our subjects were indeed experiencing AVHs by showing overall activation in speech and language areas similar to what has previously been demonstrated.[11,13] Thus, we initially performed a standard block-analysis, contrasting defined on-periods (ON-blocks) with defined off-periods (OFF-blocks).

Methods

Subjects

Structural MRI and functional MRI-BOLD data were collected from a total of 66 subjects (29 males and 37 females, mean age 38.2 (SD 13.0) years), of which 45 subjects (25 males and 20 females, mean age 37.9 (SD 13.2) years) were diagnosed with an ICD-10 or DSM-IV schizophrenia spectrum disorder.

The patients came from three collaborating projects and sites. These were University of Bergen, Norway (n = 11, 7 males, mean age 27.8 (SD 7.0) years); Plovdiv Medical University, Bulgaria (n = 13, 11 males, mean age 35.3 (SD 14.0 years) and Groningen University Medical Center, Netherlands (n = 21, 7 males, mean age 39.0 (SD 11.4) years. Symptom severity for patients was assessed with the positive and negative syndrome scale (PANSS).[29] For inclusion, a patient had to score ≥ 3 on the PANSS, P3 hallucinatory behavior item within a week of the MR scanning (mean P3 score for the whole sample 4.89 (SD 0.73)). The patients were all on second-generation antipsychotics (often clozapine). Mean PANSS positive score for the whole sample was 17.96 (SD 4.28), mean negative PANSS score was 16.91 (SD 5.18), and mean general PANSS score was 34.13 (SD 8.31). Mean total PANSS score was 69.00 (SD 15.98). The corresponding PANSS scores for each subsample and site are given in table 1. The scores shown in table 1 were subject to separate one-way ANOVAs comparing scores for each PANSS subscale between the three sites. These analyses showed no significant differences for any of the variables (right-hand column).

The patients were all on second-generation antipsychotics (often clozapine), with some patients in addition being prescribed antidepressants and/or anxiolytics. Mean antipsychotic Defined Daily Doses (DDD)[30] were 0.876 (SD 0.473), see table 1 for DDD details for the three sites. The total sample also included 21 nonclinical hallucinating individuals (4 male, mean age 44.5 (SD 13.0) years), i.e. in whom a clinical axis I or II diagnosis was ruled out using CASH and SCID-II interviews, included in the Groningen University Medical Center sample, for details see.[31]

Data Collection

Functional BOLD MR data were collected during a "symptom-capture" paradigm,[23] where subjects were instructed to press a button when a hallucinatory episode began (onset), and to press another button when the episode ended (offset). Instructions were presented visually

Table 1. Mean and standard deviation (SD) PANSS scores for P3 (hallucinatory behavior), total positive symptoms (PosTot), total negative symptoms (NegTot), total general symptoms (GenTot), and overall total (Total) scores, split for the three sites (Bergen, Groningen, Plovdiv).

	Bergen		Groningen		Plovdiv		
	Mean	SD	Mean	SD	Mean	SD	Sign.
P3	4.89	0.73	5.0	0.63	4.78	0.7	n.s.
PosTot	17.95	4.28	16.52	4.03	18.92	3.79	n.s.
NegTot	16.91	5.18	15.84	4.99	18.57	4.65	n.s.
GenTot	34.13	8.31	32.94	8.42	36.42	7.21	n.s.
Total	69.0	15.98	65.31	15.67	73.92	14.97	n.s.
DDD	0.98	0.75	0.72	0.24	0.93	0.43	n.s
Num Episodes	5.8	4.6	20.8	11.1	13.3	9.9	
Duration (s)	53	71	24	36	61	107	

Notes: DDD, Defined Daily Doses, (some missing data for the Groningen sample). One-way ANOVAs and *t*-tests for significant differences between sites. n.s. = not sign. difference. The number and duration of events is also summarized, noting variations due to differing total scan time across sites.

through LCD goggles mounted on the head-coil, in the language appropriate to the location, along with a fixation cross displayed in the middle of the visual field. A high-resolution structural T1 volume was acquired for each subject, along with additional sequences that varied between sites, not relevant to the present article. Further details on sequence parameters are presented in supplementary methods.

Data Preparation/ and Preprocessing

Visual inspection of the participants' button-press data revealed a number of redundant or ambiguous events, where a reported onset did not distinctly match to a single reported end of an episode; an operational definition detailed in supplementary methods was therefore applied to extract valid episodes; a summary of the number and duration of these episodes for each site is provided in table 1, noting that the total acquisition time differed between sites.

Functional MR data were preprocessed using standard tools from FSL 5.0.11 (FEAT pipeline), incorporating brain masking, slice-timing correction, spatial smoothing, high-pass filtering, non-linear registration to a standard-space template (via the per-subject high resolution structural image), with additional filtering of motion artifacts using the ICA-AROMA method.[32–35] Further details in supplementary methods.

Overall Block-Analysis

Preprocessed data were initially subject to a standard fMRI- BOLD block-analysis, using FSL FEAT first-level and higher-level analysis pipelines,[36] where active "episodes" were identified as the time between an onset and offset button-press which were contrasted against passive "episodes", i.e. the time outside the defined active episodes. Active and passive episodes were then treated as active and passive "blocks" in the block-analysis, following standard fMRI terminology. Mixed effects modeling (FLAME 1 + 2) was used, with clusters were thresholded nonparametrically at $z > 4.5$; FWE-corrected cluster significance thresholded at $P < .05$ and with a cluster extent > 20 voxels.

The purpose of this analysis was to allow for comparison with previous findings of overall neuronal activation *during* hallucinatory experiences, not specifically related to onsets and offsets,[11,13] and to confirm the validity of the filtered response data leading into the subsequent event-related 4D permutation analysis. Two additional variations on this analysis were performed: the first was to confirm that identified clusters were consistent for the separate collaborating sites (i.e., not dependent on particulars of the hardware and sequence), and the second to minimize possible influence of the motor-response relating to finger-movements (the button-press itself) on the findings; both of these are described in detail in supplementary material.

Event-Related Time-Course Analysis

Following the block-analysis we performed an analysis of the time-course of the BOLD response associated with on- and offset button-presses. For each button-press event demarcating the start or end of a hallucinatory episode ($t = 0$ s), windows of functional data were extracted on the range $t = [-10, +15]$ s, sampled at regular 1 s intervals. Extracted segments (for every onset- and offset-event from all subjects) were subject to permutation analysis, using a locally-developed tool (available here: https://git.app.uib.no/bergen-fmri/functional-transients) to characterize activation specific to onset- or offset-events at each time-point in the extracted window segments, voxel-wise across the entire brain volume. Reported *P*-values were modeled on a gamma approximation of the distribution after $N = 10{,}000$ permutations.[37] In addition to jointly

contrasting onset- and offset-events in the permutation analysis (yielding differential effects), the two event types were subsequently contrasted against segments extracted round random time-points (without synchronization to subjects' button-responses), representing a baseline state. A canonical double-gamma hemodynamic response function (HRF) model was fit to the extracted time-courses (optimizing magnitude and temporal origin with non-linear least squares regression) to estimate the timing of activity initiating the observed time-course.

Results

Activation Patterns During Hallucinatory Episodes

The results from the block-analysis revealed several statistically significant clusters of increased activation during hallucinatory episodes, shown in figure 1 and table 2.

These clusters included the left fronto-temporal language areas (Wernicke's area in the superior temporal gyrus, and Broca's area in the inferior frontal gyrus), regions which broadly agreed with clusters identified in previous studies, see meta-analyses.[11,13] To verify that similar patterns existed in each site's data, additional second-level contrasts were analyzed for each of the three sites separately, with the same initial cluster thresholding ($z = 4.5$) as used for the entire dataset. Although significance and (thresholded) extent varied somewhat between the groups, there was strong agreement between the location of significant clusters from the Bergen only, Bergen and Plovdiv only, and Groningen only findings and those from the entire (combined) findings. Results from the larger Groningen cohort were more statistically robust, leading to larger clusters. In this case, local maxima corresponding with discrete clusters in the full analysis are also reported. Significant activation clusters for each site are presented in supplementary table S2.

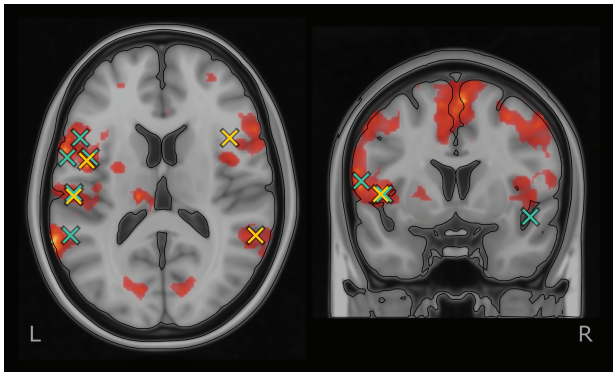

Figure 1. The results from the standard group-level block-analysis of the BOLD-fMRI data, overlaid with peak activation from the Jardri et al.[11] and Kompus et al.,[13] meta-analyses, marked respectively with a cyan (appearing darker in black and white) (Jardri) and yellow (appearing bright in black and white) (Kompus) "x", verifying the presently seen activation with activation previously repeatedly reported in the literature.

Contrasting clinical- and nonclinical voice hearers revealed overlapping activity in all areas except for the right planum temporale and lateral superior occipital cortex. The block-analysis in addition showed significant activations in the precentral gyrus, and the supplementary motor area (SMA), which most likely relate to the execution of motor- responses related to button-press activity, and not to the experience of AVHs as such. Incorporating all button-press events (including those deemed "spurious" by the operational definition) as a regressor of no interest confirmed that the primary findings were not a consequence of motor-responses related to button-press activity.

Differential Activation in the Ventromedial Prefrontal Cortex (VMPFC)

Analysis of the extracted time-windows revealed opposing time-courses for onset- versus offset events, in the ventral edge of the intersection of the paracingulate cortex, medial frontal cortex and the frontal pole, which hereafter will be referred to as ventromedial prefrontal cortex (VMPFC) (figure 2).

Panel A in figure 3 shows this differential effect (shaded blue area and corresponding solid blue line) with a significant maximum difference-peak at around $t = 2.0$ s ($\Delta = -634$ iu, $P = 4.31 \times 10^{-4}$) relative to the button-press event at $t = 0$ s. The striped blue line in Panel A shows the corresponding modeled HRF, correlating strongly with the actual data ($R^2 = 0.70$), with a lagged peak at around $t = 2.0$ s which would correspond to activation initiated around $t = -3.0$ s. This confirmed the assumption that the maximum peak-difference observed for onset- versus offset-events was initiated a few seconds in advance of the button-press at time $t = 0$ s. See figure 3 legend for further details.

Contrasting against pseudo-baseline data to determine whether differential effects (figure 3 Panel A) were driven by onset- or offset events, or both revealed a significant increase specific to onset-events with a maximum peak of 524 iu, $P = 2.9 \times 10^{-5}$ at time $t = 0$ s, corresponding with activation initiated at $t = -5.0$ s according to the HRF-model ($R^2 = 0.67$) (figure 3 Panel B), and a significant decrease specific to offset-events with a minimum peak at −979 iu, $P = 5.07 \times 10^{-5}$ at around time $t = 3$ s (figure 3 Panel C), corresponding with activation initiated around $t = -2.0$ s according to the HRF-model ($R^2 = 0.56$). The peaks for both onset- and offset-events preceded a large and significant increase in activation in the primary motor cortex in the precentral gyrus (related to the subjects' button-press), which had a maximum peak of 4843 iu, $P < 10^{-15}$ at $t = 5.0$ s relative to "baseline", see figure 3 Panel D, unrelated to the VMPFC activation. As expected, the HRF model-fit correlated strongly with activation initiated at the time the button-press was executed (derived $t = 0.1$ s, $R^2 = 0.90$).

Table 2. The Table shows clusters and local maxima obtained in the block analysis, together with corresponding activations from Jardri et al.[11] and Kompus et al..[13] Clusters in the present data are thresholded nonparametrically at $z > 4.5$; corrected cluster significance is reported, thresholded and FEW-corrected at $P > .05$ with a cluster extent > 20 voxels. MNI coord = Montreal Neurological Institute Coordinates.

				MNI coord (mm)				Jardri et al.				Kompus et al.		
								MNI coord (mm)				MNI coord (mm)		
Present study		*P*	ext	*x*	*y*	*z*			*x*	*y*	*z*	*X*	*y*	*z*
Inferior Frontal Gyrus	BA44 L	3.8×10^{-4}	43	−56	12	8	A	−50	13	5				
								Broca's convolution						
Precentral Gyrus	BA44 L			*−58*	*6*	*9*		−57	2	13				
								Precentral gyrus						
Insula	BA44 L	5.8×10^{-5}	65	−45	0.7	5.1		−44	2	3		−46	0	3
(Central Opercular Cortex)								Anterior insula				Insula		
Hippocampus							B	−26	−31	−10		−26	−31	−10
(parahippocampal gyrus)								Hippocampus/ parahippocampal gyrus				Hippocampus		
Thalamus	L	3.7×10^{-4}	21	−13	−23	7.5								
Inferior Frontal Gyrus	BA45 R	7.4×10^{-5}		56.8	19.6	4.8						42	14	13
												Inferior frontal gyrus		
Superior frontal gyrus												27	42	28
												Superior frontal gyrus		
Anterior insula							C	47	10	−10				
								Anterior insula						
Frontal Operculum								44	17	−17				
								Frontal Operculum						
Supramarginal Gyrus	BA22 L	5.3×10^{-7}	97	−65	−50	18.7	D	−56	−45	14				
(posterior division)	BA22 R							Superior and middle temporal gyri						
Postcentral Gyrus	BA2 L	6.3×10^{-13}	396	−52	−23	38.7						−51	−26	43
				−50	−29	47.4						Postcentral gyrus		
Inferior parietal lobule												31	−44	53
												Inferior parietal lobule		
Central Opercular Cortex		1.2×10^{-5}	86	−53	−23	22	E	−55	−20	14		−55	−22	15
								Supramarginalis gyrus				Superior temporal gyrus		
Middle temporal gyrus												57	−31	−11
												Middle temporal gyrus		
Middle temporal gyrus												59	−45	11
												Middle temporal gyrus		
Cerebellum lobule V												21	−45	−24
												Cerebellum lobule V		
Precentral Gyrus	BA6 L	7.75×10^{-7}	125	−56	−6	51								
				−50	−3.5	50								
				−53	−2	45								
Juxtapositional Lobule	BA6 R	4.6×10^{-13}	403	1.5	−2.8	57								
				4.7	8.7	64								
				5	6	60								

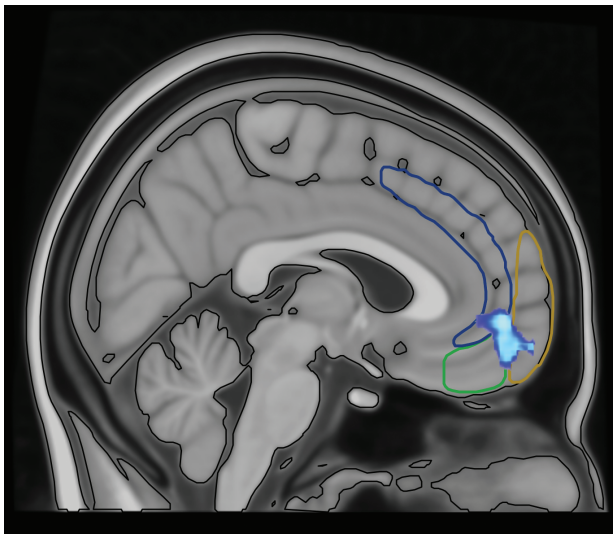

Figure 2. The anatomical localization of the VMPFC ROI (bright cyan) with MNI peak coordinates x = 8.8, y = 44.8, z = -5.64 mm from where the time-courses were extracted, in the intersection of the paracingulate cortex (demarcated in dark blue), medial inferior frontal cortex (demarcated in green) and frontal pole (demarcated in yellow).

Discussion

The block-analysis revealed that the activated ROIs showed large overlaps with the activated ROIs reported in the Jardri et al.[11] and Kompus et al.[13] meta-analyses. In this respect, the current findings replicate previous findings, a sign of robustness of the current findings, and of temporal lobe activation associated with AVHs in general. The time-series-analysis (figure 3) and the modeling of the observed time-course revealed a significant brain response initiated 2–5 s in advance of a button-press, indicating that participants experienced that they were either in or out of a hallucinatory state. This implies that the ventromedial prefrontal region is crucial in both the initiation and inhibition of hallucinatory episodes and speaks to a regulatory role, or what we would like to label a neuronal switch function for this region. Looking at figure 3, Panels B and C, it seems that the increase in activation to onset-events reached a maximum peak around 3 s before the minimum peak to offset-events. That is, the change in brain activation in advance of a conscious and aware button-press was initiated earlier when the individual subsequently experienced the disappearance of a hallucinatory episode compared to the change initiated in advance of when the individual subsequently experienced the presence of a hallucinatory episode. This could point to a difference in timing of excitatory versus inhibitory neuronal events underlying the dynamic transient fluctuations of hallucinatory episodes. Previously, Shergill et al.[27] reported activation in fronto-temporal areas 6–9 s in three patients before they signaled the presence of a voice. The timing of the findings of Shergill et al.[27] resonate with the present findings, considering the delay of the peak of the BOLD response of around 5 s after it is initiated. It should be noted, however, that these authors did not monitor activation in advance of the offset of a voice, such that we would not know from their findings if the activated regions at the onset of an episode also would be activated in advance of the offset of an episode, see also.[28,38] See however Lefebvre et al.,[39] who applied a different approach by interviewing patients afterwards, and Fovet et al.[40] who reported that frontal lobe activation predicted presence of AVHs, using statistical methods and machine learning to reliably discriminate between AVH ON and OFF periods in the fMRI time-series data.[41] Other studies have indicated aberrant functional connectivity[42] and morphological differences[43] in the VMPFC region in AVH-patients. For example, Garrison et al.[43] found that hallucinating patients had shorter paracingulate sulcus than healthy controls, and suggested that this region of the brain is tuned to "reality monitoring", i.e. the ability to judge whether a memory comes from an outer or inner source.[44] This fits with the findings by Konu et al.[45] who recently reported that the ventromedial prefrontal cortex is involved in self-generated intrusive thoughts. The present finding of changes in activation in this region before and during the onset of a hallucinatory episode then fits a role for the VMPFC in triggering self-generated mental activity. The VMPFC region also overlaps with the ventromedial part of the orbitofrontal cortex, which has been implicated in prediction errors and violation of top-down expectations and error monitoring.[46] Moreover, it has repeatedly been shown that AVHs are related to failure of predictive coding[47,48] in the sense that processing of bottom-up excitatory sensory information is distorted by aberrant top-down inhibitory processes, resulting in faulty expectations and beliefs about the environmental context. It could be argued that the VMPFC is critical for generating bottom-up/top-down expectations[44] and that AVHs occur because of an imbalance between bottom-up sensory input and top-down reality monitoring, which could be conceived as an imbalance between excitatory and inhibitory influences, the so-called E/I-imbalance model.[19,20] We now suggest that this region in the VMPFC plays a role in the balancing of excitatory and inhibitory influences, such that when it is switched on, excitatory influences are dominating, resulting in the onset of a hallucinatory state. When it is switched off, inhibitory influences dominate, resulting in the offset of a hallucinatory episode. As such the current finding could have clinical implications by opening for more targeted treatments. One option would be the use of fMRI-guided neurofeedback to train patients to control activation of this area, which in turn could help them gain control over the hallucinatory episodes they experience, either by preventing onsets or accelerating offsets. Alternatively, brain stimulation interventions, such as transcranial magnetic stimulation (TMS) and transcranial direct current stimulation

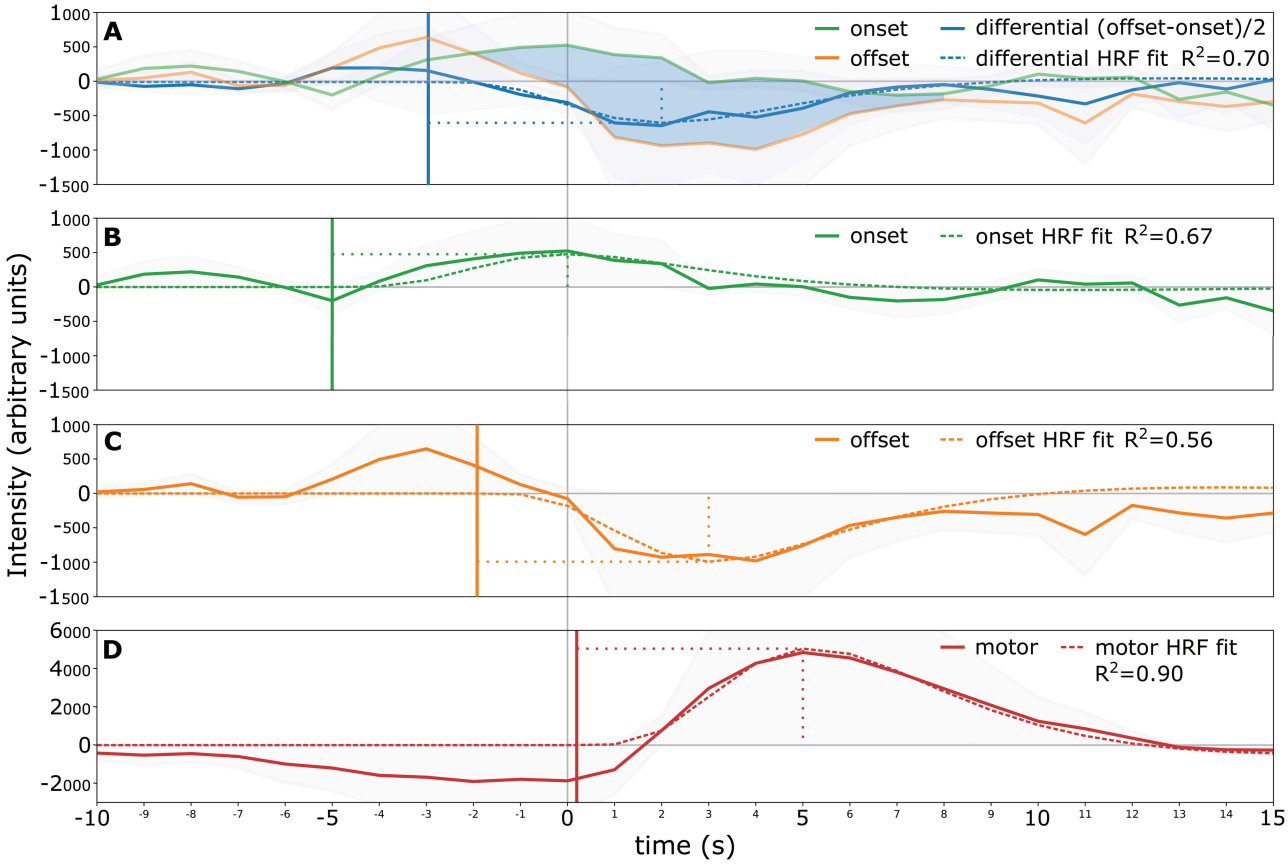

Figure 3 A-D. Mean time-courses tracked second-by-second around the onset-of-hallucination and offset-of-hallucination, with time "0" on the *x*-axis representing the point in time when the button-press occurred. The *x*-axis shows time relative to the button press (at time $t = 0$). Panel A shows differential effect (solid blue curve) from the VMPFC for onset-of-hallucinations (green curve) versus offset-of-hallucinations (orange curve). The striped blue curve shows the modeled HRF-lagged BOLD response with a lagged peak around $t = 2$ s corresponding to an expected activation initiation at around $t = -3$ s. Panel B shows onset-of-hallucination episodes (solid green curve, with modeled HRF fit as a striped green curve) from the VMPFC compared time-courses extracted from random periods within the session (striped green curve). Panel C shows corresponding offset-of-hallucination episodes (solid orange curve, with modeled HRF-fit as a striped orange curve) compared with time-courses extracted from random periods within the session. Panel D shows activation from a region in the pre-central motor cortex (solid red curve, with modeled HRF fit as a striped red curve), compared with time-courses extracted from random periods within the session. Note different *y*-axis scale in this panel.

(tDCS) may also be used to targeting the VMPFC to stop the onset or accelerate the offset of an AVH-episode.

Strengths and Limitations

An obvious strength is that we could gather a satisfactory number of participants and MR-scans who were able to signal their hallucinations, which enabled us to detect short-lasting neuronal phenomena like the switching between hallucinatory and non-hallucinatory states. Another strength is the multicentric nature of the data, and the inclusion of nonclinical AVH individuals, which adds a natural variability to the data, resulting in increased generalizability. Although we used data from three different scanners (and in fact from three different countries), and multiple scanner input may have introduced extra variability, the fact that we nevertheless could find significant activation changes points to a rather robust effect. A potential limitation of the results could be anticipatory attention focus on motor-responses, cf.,[49] which otherwise could affect the observed activation. As seen in the lower panel (D) of figure 3, this is probably not the case, since there was a clear peak around 5 s post-response obtained from the pre-central motor-cortex region on the left side, and with 5-6 times higher response amplitude as that obtained from the VMPFC region. As seen in the HRF-modeled time-courses in figure 3, this is what one would expect considering the lag of the hemodynamic response relative to a neuronal event.[50] Adding to this is the possibility that the activation seen in the VMPFC is related to response anticipation and motor planning. There is a rather extensive brain imaging literature on planning of simple motor responses, see examples in.[51–53] For example, Seghezzi et al.[52] performed a meta-analytical comparison of imaging studies on motor intentionality and their findings did not include the VMPFC. Similar results were reported by Lee et al.[53] using EEG, and by Yeom et al.[51] using MEG, see also.[54] It is therefore unlikely that the activation found in the VMPFC in the present study was caused by the intention to move the fingers to press the response button. This

argument is strengthened by the fact that this region was differentially activated to onset versus offset responses although the intention to press a button would be the same in these two conditions. As mentioned in the Results section, the observed supplementary motor area (SMA) activation was most likely related to the execution of motor- responses related to button-press activity, and not to the experience of AVHs as such. This conclusion is corroborated by the finding of Linden et al.[55] that SMA activation did not precede temporal lobe activation during hallucinations as it did during verbal imagery. Another potential limitation is the possible confounding of finger- and head-movements in advance of on- and offsets. A supplementary analysis explored these aspects and revealed only weak correlation with motion parameters (reduced after filtering with ICA-AROMA) and only a small number of residual components correlating moderately with the button-press events (see supplementary figure S2). The comparison of the BOLD time-courses in the permutation analysis would also rule out confounding by motor-activity since the motor-responses would be similar for onset and offset events. There was a difference with right versus left hand presses for onsets and offsets, but this difference should not have affected activations along the midline, including the current VMPFC activation, and regressing out motor (button-press) responses did also not confound the results.

Conclusions

In conclusion, we identified an area in the VMPFC, which showed increased activation preceding the start of a hallucinatory episode and decreased activation preceding the end of a hallucinatory episode. This hallucinatory switch could be a target for new interventions such as fMRI-guided neurofeedback or brain stimulation.

Supplementary Material

Supplementary material is available at *Schizophrenia Bulletin*.

Funding

This work was supported from the European Research Council (ERC) #249516, Western Norway Health Authorities #912045, Medical University of Plovdiv, Bulgaria, and ZonMW (Dutch Medical Research Council) # 91213009. Co-authors Kenneth Hugdahl, Alexander Craven and Lars Ersland own shares in the company NordicNeurolab Inc, which produced additional equipment used for the fMRI image acquisitions. The authors declare no conflict of interest.

Ethics approval

The study was approved by the local ethics committees at each site, and had a European Research Council Ethics Approval (ERCEA 2016-439428). Transfer of data between the sites and re-analysis at the Bergen University was approved by the ethical committees at each site, and further confirmed by the Regional Committee for Medical Research Ethics in Western Norway (REK-Vest #2017/933).

Acknowledgments

The authors want to acknowledge the contribution by MR-technicians and participants for making the study possible. Special thanks to Marco Hirnstein, University of Bergen, Norway for constructive comments on an earlier version of the manuscript.

References

1. Andreasen NC, Olsen S. Negative v positive schizophrenia. Definition and validation. *Arch Gen Psychiatry.* 1982;39:789–794.
2. Sartorius N, Jablensky A, Korten A, et al. Early manifestations and first-contact incidence of schizophrenia in different cultures. A preliminary report on the initial evaluation phase of the WHO collaborative study on determinants of outcome of severe mental disorders. *Psychol Med.* 1986;16:909–928.
3. Waters F, Badcock J, Michie P, Maybery M. Auditory hallucinations in schizophrenia: Intrusive thoughts and forgotten memories. *Cogn Neuropsychiatry* 2006;11:65–83.
4. Aleman A, Larøi F. Hallucinations: The Science of Idiosyncratic Perception. Washington, DC: American Psychological Association; 2008.
5. Ford JM, Dierks T, Fisher DJ, et al. Neurophysiological studies of auditory verbal hallucinations. *Schizophr Bull.* 2012;38:715–723.
6. Hugdahl K, Sommer IE. Auditory verbal hallucinations in schizophrenia from a levels of explanation perspective. *Schizophr Bull.* 2018;44:234–241.
7. Ford JM, Morris SE, Hoffman RE, et al. Studying hallucinations within the NIMH RDoC framework. *Schizophr Bull.* 2014;40:295–304.
8. Zhuo C, Zhou C, Lin X, et al. Common and distinct global functional connectivity density alterations in drug-naïve patients with first-episode major depressive disorder with and without auditory verbal hallucination. *Prog Neuropsychopharmacol Biol Psychiatry.* 2020;96:1–6.
9. Garcia-Ptacek S, Kramberger MG. Parkinson disease and dementia. *J Geriatr Psychiatry Neurol.* 2016;29:261–270.
10. Korsnes MS, Hugdahl K, Nygard M, Bjornaes H. An fMRI study of auditory hallucinations in patients with epilepsy. *Epilepsia* 2010;51:610–617.
11. Jardri R, Pouchet A, Pins D, Thomas P. Cortical activations during auditory verbal hallucinations in schizophrenia: A coordinate-based meta-analysis. *Am J Psychiatry.* 2011;168:73–81.
12. Lavigne KM, Rapin LA, Metzak PD, et al. Left-dominant temporal-frontal hypercoupling in schizophrenia patients with hallucinations during speech perception. *Schizophr Bull.* 2015;41:259–267.
13. Kompus K, Westerhausen R, Hugdahl K. The "paradoxical" engagement of the primary auditory cortex in patients with auditory verbal hallucinations: A meta-analysis of functional neuroimaging studies. *Neuropsychologia* 2011;49:3361–3369.

14. Aleman A, Vercammen A. Functional neuroimaging of hallucinations. In: Blom JD, Sommer IEC, eds. *Hallucinations—Research and practice*. London: Springer Science & Business Media; 2012:267-281.
15. McCarthy-Jones S, Trauer T, Mackinnon A, Sims E, Thomas N, Copolov DL. A new phenomenological survey of auditory hallucinations: Evidence for subtypes and implications for theory and practice. *Schizophr Bull* 2014;40:231-5.
16. Allen P, Larøi F, McGuire PK, Aleman A. The hallucinating brain: A review of structural and functional neuroimaging studies of hallucinations. *Neurosci Biobehav Rev.* 2008;32:175–191.
17. Zmigrod L, Garrison JR, Carr J, Simons JS. The neural mechanisms of hallucinations: A quantitative meta-analysis of neuroimaging studies. *Neurosci Biobehav Rev.* 2016;69:113–123.
18. Ćurčić-Blake B, Ford JM, Hubl D, et al. Interaction of language, auditory and memory brain networks in auditory verbal hallucinations. *Prog Neurobiol.* 2017;148:1–20.
19. Steinmann S, Leicht G, Mulert C. The interhemispheric miscommunication theory of auditory verbal hallucinations in schizophrenia. *Int J Psychophysiol.* 2019;145:83–90.
20. Jardri R, Hugdahl K, Hughes M, et al. Are hallucinations due to an imbalance between excitatory and inhibitory influences on the brain? *Schizophr Bull.* 2016;42:1124–1134.
21. Allen P, Sommer IE, Jardri R, Eysenck MW, Hugdahl K. Extrinsic and default mode networks in psychiatric conditions: Relationship to excitatory-inhibitory transmitter balance and early trauma. *Neurosci Biobehav Rev.* 2019;99:90–100.
22. Moonen CTW, Bandettini PA, eds. *Functional MRI*. Berlin, Germany: Springer-Verlag; 1999.
23. Sommer IE, Diederen KM, Blom J-D, et al. Auditory verbal hallucinations predominantly activate the right inferior frontal area. *Brain* 2008;131:3169–3177.
24. Dierks T, Linden DEJ, Jandl M, et al. Activation of Heschl's gyrus during auditory hallucinations. *Neuron* 1999;22:615–621.
25. van de Ven VG, Formisano E, Röder CH, et al. The spatiotemporal pattern of auditory cortical responses during verbal hallucinations. *Neuroimage* 2005;27:644–655.
26. Shergill SS, Brammer MJ, Williams SC, Murray RM, McGuire PK. Mapping auditory hallucinations in schizophrenia using functional magnetic resonance imaging. *Arch Gen Psychiatry.* 2000;57:1033–1038.
27. Shergill SS, Brammer M, Amaro E, Williams S, Murray R, Mcguire P. Temporal course of auditory hallucinations. *Br J Psychiatry.* 2004;185:516–517.
28. Hoffman RE, Anderson AW, Varanko M, Gore JC, Hampson M. Time course of regional brain activation associated with onset of auditory/verbal hallucinations. *Br J Psychiatry.* 2008;193:424–425.
29. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull.* 1987;13:261–276.
30. Leucht S, Samara M, Heres S, Davis JM. Dose equivalents for antipsychotic drugs: The DDD Method. *Schizophr Bull.* 2016;42(Suppl 1):S90–S94.
31. Sommer IE, Daalman K, Rietkerk T, et al. Healthy individuals with auditory verbal hallucinations; Who are they? Psychiatric assessments of a selected sample of 103 subjects. *Schizophr Bull.* 2010;36:633–641.
32. Jenkinson M, Smith SA. global optimisation method for robust affine registration of brain images. *Med Image Anal.* 2001;5:143–156.
33. Woolrich MW, Ripley BD, Brady M, Smith SM. Temporal autocorrelation in univariate linear modeling of fMRI data. *Neuroimage* 2001;14:1370–1386.
34. Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM. FSL. *Neuroimage* 2012;62:782–790.
35. Pruim RHR, Mennes M, van Rooij D, Llera A, Buitelaar JK, Beckmann CF. ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data. *Neuroimage* 2015;112:267–277.
36. Woolrich MW, Behrens TEJ, Beckmann CF, Jenkinson M, Smith SM. Multilevel linear modelling for fMRI group analysis using Bayesian inference. *Neuroimage* 2004;21:1732–1747.
37. Winkler AM, Ridgway GR, Douaud G, Nichols TE, Smith SM. Faster permutation inference in brain imaging. *NeuroImage* 2016;141:502–516.
38. Diederen KM, Neggers SF, Daalman K, et al. Deactivation of the parahippocampal gyrus preceding auditory hallucinations in schizophrenia. *Am J Psychiatry.* 2010; 167:427–435.
39. Lefebvre S, Demeulemeester M, Leroy A, et al. Network dynamics during the different stages of hallucinations in schizophrenia. *Hum Brain Mapp.* 2016;37:2571–2586.
40. Fovet T, Yger P, Lopes R, et al. Decoding activity in Broca's area predicts the occurrence of auditory hallucinations across subjects. *Biol Psychiatry.* 2022;91(2):194–201.
41. de Pierrefeu A, Fovet T, Hadj-Selem F, et al. Prediction of activation patterns preceding hallucinations in patients with schizophrenia using machine learning with structured sparsity. *Hum Brain Mapp.* 2018;39(4):1777–1788.
42. Alonso-Solís A, Vives-Gilabert Y, Grasa E, et al. Resting-state functional connectivity alterations in the default network of schizophrenia patients with persistent auditory verbal hallucinations. *Schizophr Res.* 2015;161:261–268.
43. Garrison JR, Fernyhough C, McCarthy-Jones S, Simons JS, Sommer IE. Paracingulate sulcus morphology and hallucinations in clinical and nonclinical groups. *Schizophr Bull.* 2019;45:733–741.
44. Buda M, Fornito A, Bergström ZM, Simons JS. A specific brain structural basis for individual differences in reality monitoring. *J Neurosci* 2011;31:14308-13.
45. Konu D, Turnbull A, Karapanagiotidis T, et al. A role for ventromedial prefrontal cortex in self-generated episodic social cognition. *Neuroimage* 2020;218:116977.
46. Rudebeck PH, Rich EL. Orbitofrontal cortex. *Curr Biol.* 2018;28:R1083–R1088.
47. McCleery A, Wynn JK, Mathalon DH, Roach BJ, Green MF. Hallucinations, neuroplasticity, and prediction errors in schizophrenia. *Scand J Psychol.* 2018;59: 41–48.
48. PowersARIII, Kelley M, Corlett PR. Hallucinations as top-down effects on perception. *Biol Psychiatry Cogn Neurosci Neuroimaging* 2016;1:393–400.
49. van Lutterveld R, Diederen KM, Schutte M, Bakker R, Zandbelt B, Sommer IE. Brain correlates of auditory hallucinations: Stimulus detection is a potential confounder. *Schizophr Res.* 2013;150:319–320.
50. Logothetis NK, Pfeuffer J. On the nature of the BOLD fMRI contrast mechanism. *Magn Reson Imaging.* 2004;22:1517–1531.
51. Yeom HG, Kim JS, Chung CK. Brain mechanisms in motor control during reaching movements: Transition of functional connectivity according to movement states. *Sci Rep.* 2020;10:567.

52. Seghezzi S, Zirone E, Paulesu E, Zapparoli L. The brain in (willed) action: a meta-analytical comparison of imaging studies on motor intentionality and sense of agency. *Front Psychol.* 2019;10:804.
53. Lee M, Yoon JG, Lee SW. Predicting motor imagery performance from resting-state EEG using dynamic causal modeling. *Front Hum Neurosci.* 2020;14:321.
54. Sanderson RW, Fang Q, Curatolo A, et al. Camera-based optical palpation. *Sci Rep.* 2020;10:15951.
55. Linden DE, Thornton K, Kuswanto CN, Johnston SJ, van de Ven V, Jackson MC. The brain's voices: comparing nonclinical auditory hallucinations and imagery. *Cereb Cortex.* 2011;21(2):330–337.