

Neurofeedback-based enhancement of single-trial auditory evoked potentials: treatment of auditory verbal hallucinations in schizophrenia

Kathryn Rieger^{1,2}, Marie-Helene Rarra¹, Laura Diaz Hernandez¹, Daniela Hubl¹, Thomas Koenig^{1,2}

¹Translational Research Center, University Hospital of Psychiatry, University of Bern, Bern, Switzerland

²Center for Cognition, Learning and Memory, University of Bern, Bern, Switzerland

Corresponding author:

Prof. Thomas Koenig
Translational Research Center
University Hospital of Psychiatry
Bolligenstrasse 111
CH-3000 Bern 60, Switzerland
Phone: +41 31 9309369
E-mail: thomas.koenig@upd.unibe.ch

Abstract

Auditory verbal hallucinations depend on a broad neurobiological network ranging from the auditory system to language as well as memory-related processes. As part of this, the auditory N100 ERP component is attenuated in patients with schizophrenia, with stronger attenuation occurring during auditory verbal hallucinations. Changes in the N100 component presumably reflect disturbed responsiveness of the auditory system towards external stimuli in schizophrenia. With this premise, we investigated the therapeutic utility of neurofeedback training to modulate the auditory-evoked N100 component in patients with schizophrenia and associated auditory verbal hallucinations. Ten patients completed electroencephalography neurofeedback training for modulation of N100 (treatment condition) or another unrelated component, P200 (control condition).

On a behavioral level, only the control group showed a tendency for symptom improvement in the PANSS total score in a pre-/post comparison ($t_{(4)} = 2.71$, $p = 0.054$), however, no significant differences were found in specific hallucination related symptoms ($t_{(7)} = -0.53$, $p = 0.62$). There was no significant overall effect of neurofeedback training on ERP components in our paradigm; however, we were able to identify different learning patterns, and found a correlation between learning and improvement in AVH symptoms across training sessions ($r = 0.664$, $n = 9$, $p = 0.05$). This effect results, with cautious interpretation due to the small sample size, primarily from the treatment group ($r = 0.97$, $n = 4$, $p = 0.03$). In particular, a within-session learning parameter showed utility for predicting symptom improvement with neurofeedback training.

In conclusion, patients with schizophrenia and associated auditory verbal hallucinations who exhibit a learning pattern more characterized by within-session aptitude may benefit from electroencephalography neurofeedback.

Furthermore, independent of the training group, a significant spatial pre-post difference was found in the event-related component P200 ($p = 0.04$).

Keywords

Auditory verbal hallucinations, event-related potentials, N100, neurofeedback, P200, schizophrenia

Introduction

Auditory verbal hallucinations (AVH) are a common symptom of schizophrenia that is associated with significant psychological strain. These hallucinations can occur on a broad spectrum, with voices heard in varying volumes from one or several people with comprehensible or indistinguishable voices and friendly or hostile intents. AVHs can manifest occasionally or continuously in patients. Together, these features reflect substantial heterogeneity in AVHs across patients.

Research in the last two decades indicates that AVHs arise from the abnormal and complex neurobiological interaction of language, auditory, and memory brain networks¹. Evidence specifically implicates the involvement of basic cortical areas such as the primary auditory cortex (PAC)², secondary auditory cortex, and premotor cortex; complex cortical structures including the prefrontal cortex and anterior cingulate cortex (ACC)³; and subcortical and cerebellar regions (for an overview, see^{4,5}). On this premise, it is not surprising that 25–30% of patients with schizophrenia are resistant to conventional treatment with antipsychotic medication⁶. Alternative approaches such as psychotherapy⁷ and non-invasive brain stimulation with transient direct current stimulation (tDCS)^{8,9} or repetitive transcranial magnetic stimulation (rTMS)^{10,11} have been implemented in patients with refractory schizophrenia with limited success. Accordingly, it has been hypothesized that there are heterogeneous biological subtypes of dysfunction that underlie phenotypes of AVH.

Neurofeedback (NFB) is a therapeutic modality in which subjects learn to self-regulate brain activity through operant conditioning¹². NFB has been applied as an alternative therapeutic approach in several neuropsychiatric disorders including attention deficit hyperactivity disorder¹³⁻¹⁶, epilepsy¹⁷⁻¹⁹, and autism²⁰. In the context of schizophrenia, two early studies showed the feasibility of NFB for training the self-regulation of slow cortical potentials in patients^{21,22}. The potential therapeutic use of NFB training for schizophrenia-associated AVHs was first proposed by McCarthy-Jones²³. To this end, two recent studies reported the ability of patients with schizophrenia to control activity in the ACC using real-time functional magnetic resonance imaging NFB^{3,24}, the latter of which reported improvements in some aspects of AVH symptoms. Other more cost-effective tools for NFB

exist, including electroencephalography (EEG)-NFB; however, no study to date has explored EEG-NFB for the treatment of AVHs in patients with schizophrenia.

Motivated by our own findings identifying the essential involvement of the PAC in AVH pathogenesis^{2, 25}, we examined the therapeutic utility of EEG-NFB for schizophrenia-associated AVHs using the N100 event-related potential (ERP) component as a training target. The N100 peaks between 80 and 120 ms after an auditory stimulus²⁶, and its source has been localized in the PAC and secondary auditory cortex, ACC, and other parts of the temporal lobe and insula²⁷⁻²⁹. The amplitude of N100 is generally attenuated in patients with schizophrenia^{26, 30-32} and is further attenuated during hallucinations²⁵. Therefore, we hypothesized that decreases in the amplitude of N100 might be an index of brain network dysfunction related to AVHs in patients with schizophrenia.

A comparable background of the neurophysiology for AVH can be found in tinnitus patients. Those patients also showed a reduction of the ERP component N100 generated in the PAC, and were able to actively manipulate the amount of phase-locking onto an auditory stimulus by varying selective attention to competing auditory stimuli^{33, 34}. Further, rtfMRI NFB of the PAC may reduce symptoms in patients with tinnitus³⁵.

On this premise, we tested the feasibility of N100 amplitude modulation training with EEG-NFB for improving AVH symptoms in schizophrenia. To control for non-specific effects, we included a second group that was trained to modulate another ERP component, P200, as a control condition. No alterations in the ERP P200 were reported in schizophrenia³¹.

Further, we investigated the effects of different learning parameters following a design previously implemented by Diaz and colleagues³⁶. Differentiating these learning parameters allowed us not only to analyze a time dependent change as function of being on task, classically labelled as “learning”, but also permitted the investigation of a time effect independent of being on task (labelled as “carry-over”) and an effect of being on task independent of time (labelled as “aptitude”). A carry-over effect visualizes the ability to passively learn to change a certain brain state. Aptitude quantifies the general ability of a subject to voluntarily modulate a certain brain state and could act as a predictor for a positive therapeutic effect of EEG-NFB training on AVH symptoms.

Material and methods

Subjects

We screened a total of 170 inpatients and outpatients with schizophrenia who visited the University Hospital of Psychiatry in Bern over the two-year study recruiting period. Of these, twelve patients met the inclusion criteria and agreed to participate in our double-blind study. Two patients dropped out of the study after the psychopathological assessment, such that a total of ten patients (four women and six men; mean age, 36.6 ± 14.6 years; age range, 19–61 years) were enrolled. One patient stopped after six, another after 14 sessions out of 16 sessions. Both had been randomly assigned to the control group (P200). Eight patients were right-handed, one was left-handed, and one was ambidextrous according to the handedness questionnaire³⁷. Patients were between the ages of 18 and 65 years, able to participate in continuous NFB training over a two-week period, and had a verified diagnosis of schizophrenia (ICD-10: F20) or a schizoaffective disorder (ICD-10: F25). All patients were actively experiencing AVHs in their current phase of disease. Patients were excluded if they had another active psychiatric disease, neurological disease, problems with hearing or non-corrected vision, or were pregnant.

Patients were randomly assigned to either the treatment group (N100 training) or the control group (P200 training) (Table 1) using a stratified randomization. All patients were fully informed about the nature and purpose of the study and provided written informed consent for participation. All procedures were approved by a local ethics committee (Kantonale Ethikkommission Bern, reference number: 193/13) and conducted in accordance with the most recent version of the Declaration of Helsinki.

Design and questionnaires

Figure 1 summarizes the experimental procedure. The experiment included two pre-sessions (1), eight NFB double sessions (2), and one post-session (3).

Figure 1 about here

In the first pre-session (1), patients were fully informed about the experimental procedure and inclusion and exclusion criteria were evaluated. An interviewer collected demographic information and administered a set of questionnaires in a semi-structured interview format. Questionnaires included a) the Edinburgh Handedness Inventory ³⁷ to assess handedness, b) the Positive and Negative Syndrome Scale (PANSS) ³⁸ to assess psychotic symptoms, and c) the hallucination subscale of the Psychotic Rating Scale (PsyRats) ³⁹ to assess different dimensions of AVH. Additionally, patients completed a short battery of cognitive tests including a) the Hopkins Verbal Learning Test-Revised (HVLTR) of the MATRICS Consensus Cognitive Battery ^{40, 41} to assess verbal learning, b) the Neuropsychological Assessment Battery of the MATRIC Consensus Cognitive Battery (NAB-Mazes) ^{40, 41} to screen executive function, and c) the D2 Test of Attention ⁴² to assess selective and sustained attention. The second pre-session (1) included a hearing test to confirm normal hearing ability (Diatec Screening Audiometer, AS 608) and a series of EEG recordings to construct filters used for later analyses (see *EEG recording*).

For NFB training (2), patients completed eight double sessions over a two-week period with four training days per week. Each single-session lasted approximately 22 minutes, followed by a short break and a subsequent session of the same duration. After each single-session, patients verbally reported the strategies they used during training. After each double session, patients completed the Hallucination Change Scale (HCS) modified from Hoffmann et al. ⁴³ to assess overall changes in hallucination severity (independent of hallucination dimensions). HCS scores ranged from -10 (voices completely disappeared) to +10 (voices substantially increased). The first NFB session served as a baseline for subsequent sessions and thus was assigned a HCS score of 0 by default.

The post-session (3) was conducted on a separate day after completion of the eight double sessions and included administration of a) the PANSS to assess symptom severity of positive and negative symptoms and b) the hallucination subscale of the PsyRats ³⁹ to assess dimensions of AVH after the last training session. The NFB training was identical to our previous study in healthy controls ⁴⁴ and is briefly summarized as follows:

Patients were comfortably seated in a sound-shielded EEG cabin at a distance of 1.30 m from a computer monitor. All recordings were video-monitored to control for body movements. For EEG recordings, an ActiCap system with 32 channel-active electrodes (layout based on the 10–10 system with reference at FCz and ground at AFz) (Brain

Products GmbH, Germany) was used. Impedances were maintained below 20 k Ω . Signal were amplified (BrainAmp 32-channel amplifier, BrainProducts GmbH) band pass-filtered (0.3–70 Hz) and sampled at 500 Hz.

ERP feature extraction

In accordance with Hubl et al.²⁵, a simple series of 1000-Hz sinusoidal tones with a duration of 70 ms (rise 10 ms, plateau 50 ms, fall 10 ms) was used to evoke the target ERP component (N100 or P200). Beep tones were presented with an inter-stimulus interval of 1430 ± 140 ms and delivered by a standard hi-fi audio amplifier (Technics V300 Mark 2, Panasonic, Japan) .

Pre-session recording

Figure 2 shows a schematic of the recording and analysis procedure. In the pre-session, 3 EEGs were recorded: First a) an 8-min resting-state EEG with eyes-open/eyes-closed (alternating every 2 min), followed by b) recording of voluntary eye movements (blinking, left-right movement, and up-down movement) to allow the offline construction of c) a spatial filter by means of an independent component analysis for later online removal of ocular movements, and (3) a 4-min recording during exposure to the beep tones later used in NFB training (Figure 2, item d).

Figure 2 about here

Individual temporal and spatial filters (Figure 2, items e-f) were constructed from the pre-session EEG data with the stimulation to extract ERP amplitudes on a single-trial level as follows: first, the EEG data recorded during stimulation was band pass-filtered from 1–30 Hz, recomputed to the average reference, and segmented from 0–500 ms post-stimulus and baseline-corrected. A wavelet transformation using real Gabor-functions was performed on single-trial ERPs. To obtain the most sensitive frequency filter and latency, 1-sample t-tests were computed across trials for all electrodes, time points, and wavelet layers. The absolute values of these t-values were averaged across channels and used as a global index of the overall signal-to-noise ratio across time, frequency and space. For the detection of single-trial N100 and P200 amplitudes, the optimal individual latency and frequency layer were identified at the peak of these average t-values in the corresponding time range. Spatial filters were computed as

average maps across trials at the optimal frequency and latency (Figure 2, item g). Single-trial N100 and P200 amplitudes were thus defined as the weighted sum of all electrodes of the wavelet-filtered EEG at the individually defined latency, where the weights were defined by the spatial filter.

Online analysis of NFB training

Online single-trial extraction of the N100 amplitude was performed using build-in and in-house developed plug-ins for RecView (Brain Products GmbH). The raw data was filtered online with the previously constructed spatial filter for ocular movements and recomputed to the average reference. Then, a spatial filter and a temporal wavelet filter, defined based on the individual optimal frequency and latency, were applied to the resultant EEG data to extract the N100 ERP component on a single-trial level (Figure 2, item h). Visual feedback was provided during NFB training based on this information (Figure 2, item i).

Visualization of the feedback was provided using PsychoPy⁴⁵, depicted a balloon rising in accordance with the N100 amplitude. The following rule was used to update the position of the balloon:

$$y_t = y_{t-1} + (s - y_{t-1}) * k$$

where y_t is the value fed back to the subject at trial t , s is the value of the single-trial quantifier of the N100 amplitude at trial t , and k is a constant (0.2 in the current experiment).

NFB training protocol

Each NFB training session consisted of 3 different conditions: a passive condition and 2 active conditions (training and transfer). Each session included 8 2.5-min blocks with the following structure: passive condition, 4 training blocks, passive condition, transfer block, passive condition. During all blocks, beep tones were used to evoke the target components. For the first passive condition block, patients were instructed as follows: “Try to relax and listen to the beep tone while focusing the screen without thinking about anything in particular.” The visual feedback symbol was a static frame (i.e., no feedback). Based on the values of the target component amplitude in the first passive condition block, a threshold was set that predicted a positive feedback 60% of the time. Patients assigned to the treatment group received feedback on the N100 component, whereas patients assigned to the control group

received feedback on the P200 component. The consecutive 4 training blocks were accompanied by the following instruction: “Try to make the balloon fly as high as possible.” Patients were instructed to use mental strategies to regulate the height of the balloon. These strategies could be changed during the blocks and/or across repetitions. The last condition (“transfer”) was between 2 passive conditions; during this condition, patients did not receive any feedback while fixating the screen (i.e., there was no balloon displayed on the screen). The instruction was as follows: “Try to reproduce the same state as in the training blocks where the balloon was flying high by using the same mental strategies.” After each block, patients were allowed to take an individually defined short break before pressing a button to continue on to the next block.

Statistical analysis

Table 2 depicts an overview of the main analysis that were performed. The main goal of the analyses was to quantify systematic improvements of the clinical symptoms through NFB training, and NFB related changes of the ERP components. In addition, we were interested in knowing if there were systematic relationships between the NFB learning patterns, pre-training psychopathology, and changes in hallucinations. Given the complexity of the data, these analyses included a part that reduced the dimensionality of the psychopathology and neurophysiology data using principal component analyses (Table 3), followed by correlational analyses linking these reduced set of features. All analyses used a threshold of $p < 0.05$ for statistical significance. All statistical analyses were performed using the Statistical Package for the Social Sciences version 24 (SPSS, IBM). Linear mixed models were estimated in R⁴⁶.

Demographics, pre-and post-training group differences in clinical scores

To test whether the NFB training had a systematic effect on schizophrenia psychopathology (Table 2, I), PANSS total score and PsyRats scores of the hallucination subscale were compared for each group between pre- and post-session using paired t-tests and Wilcoxon rank-sum tests, respectively.

Table 2 about here

Pre- post differences in hallucinations

The primary outcome variables of our study were changes in clinical features after the training compared to the pre-training (Table 2, II). Thus, the HCS of the second-to-last session was subtracted from that of the first session to yield an HCS difference score; a positive difference score indicated a decrease in hallucinations. Data from the second-to-last session was used due to the dropout of a subject in the last session. To test whether NFB training had a systematic effect on AVH severity, the HCS difference score was compared to 0 across all subjects. Similarly, differences in the pre-post changes in PANSS total score and PsyRats scores of the hallucination subscale were computed and tested against zero using one-sample t-tests.

Group effects in pre- post differences in hallucinations

To test whether the training type (N100 vs P200 feedback) systematically affected the changes in hallucinations, the HCS pre-post difference score were compared between training groups using an independent-samples t-test (Table 2, III). Similar tests were computed for group differences in pre-post changes in PANSS total score and PsyRats scores of the hallucination subscale.

Parametrization of individual NFB training patterns

For the parameterization of the NFB training pattern (Table 2, IV), the neurophysiological data was individually modeled across and within sessions. A description of the different learning parameters across and within sessions is shown in table 3, and briefly reviewed here: For across- and within-session analyses, we distinguished three different learning parameters in accordance with the methods of Diaz Hernandez et al.³⁶: First, a basic ability to increase the AEP component as a function of task, henceforth referred to as “aptitude” (training aptitude and transfer aptitude), second voluntary learning, i.e. a time dependent change occurring selectively during the task, and third involuntary learning, i.e. a task independent change across time. Thus, voluntary learning describes an increase over time in the component amplitude in one of the active conditions (training or transfer) compared to the passive condition as a baseline, and is henceforth referred to as either “training learning” or “transfer learning.” Involuntary learning describes amplitude increases over time in the passive condition either within or across sessions and is henceforth referred to as “carry-over.” To assess these different learning parameters in each subject, we used a linear regression model that included a random factor defined as time (i.e., session for the across-session

analysis and repetition for the within-session analysis), a random factor defined as condition (passive or active: training or transfer), and a random interaction factor of time and condition, following a design that had been previously implemented by Rieger and colleagues⁴⁴. The use of a linear mixed effect model allowed handling incomplete data sets. Therefore, all data sets with at least six sessions were included in the analyses of the neurophysiological data. The obtained β -values were tested for significance across subjects using t-statistics. To reduce the number of tests in the further analyses, the set of individual learning parameters for the within and across session effects were submitted to a principal component analysis (PCA). From this PCA, individual scores of the first component were extracted and used as reduced quantifiers of NFB training patterns (later referred to as *learning factor*).

Group comparisons of NFB training patterns

In order to compare the two patient groups, the linear mixed models described above included group as a categorical predictor. This allowed us to test for group specific differences in the above outline NFB training patterns (Table 2, V).

Table 3 about here

Relationship of hallucination changes and NFB learning type

It is quite possible that individual NFB learning patterns have an effect on the clinical outcome of the training. To test this hypothesis, we correlated the individual pre- post differences in hallucinations with the individual factor score summarizing the NFB learning pattern (the learning factor) using a Pearson correlation coefficient (Table 2, VI).

Extraction of psychopathological characteristics

The individual psychopathology may be an important, but high-dimensional moderator in our analyses. To reduce problems of multiple testing, we therefore reduced the dimensionality of the psychopathology data, again using a PCA, and including the data of the individual sub-scores of the PsyRats, PANSS positive and negative scores, the

Calgary Depression Scale Schizophrenia (CDSS) ⁴⁷ and the wellbeing sub-score of the CORE ⁴⁸. From this PCA, individual scores of the first component were extracted and used as reduced quantifiers of individual psychopathology (Table 2, VII; Table 3).

Relationship of pre-training psychopathology and NFB learning type

To test whether the pre-training psychopathology was predictive for NFB learning success, we correlated the individual factor scores for the pre-training psychopathology with the individual factor scores obtained for the NFB training patterns, i.e. with the learning factor (Table 2, VIII).

Pre-post differences in ERP topography and source estimates

Finally, to assess whether NFB training caused changes in ERP topography and source estimates, averaged ERPs during the first block of passive listening were compared between first and last training sessions (Table 2, IX). To localize brain regions affected by the training, source estimates of the ERPs were computed as follows: ERPs were band pass-filtered from 1.5–15 Hz, recomputed to the average reference, and corrected for eye-movement artifacts with an independent component analysis. The time windows of the N100 and P200 components were defined using a microstate analysis ⁴⁹ that parsed data into periods of quasi-stable field configurations, thereby collapsing data across time periods with presumably stable source configuration. Within the N100 and P200 microstates, we compared the amplitude-normalized spatial configurations of the ERP fields using topographical analyses of variance (TANOVAs) ⁵⁰ and the mean global field power (GFP), which quantifies the overall potential variance across electrodes ⁵¹ between pre- and post-training. Significant TANOVA results were further analyzed by statistical mapping of sLORETA voxel-wise inverse solutions to estimate the sources accounting for training effects. sLORETA inverse solutions were computed using sLORETA software ⁵² using a boundary elements head model derived from the MNI 152 template with a total of 6239 voxels (voxel size, $5 \times 5 \times 5 \text{ mm}^3$) covering the gray matter of the brain. The regularization parameter SNR was set to 100. Paired t-statistics across subjects were used to assess significant local differences. Given that the overall null hypothesis was rejected by the previous TANOVA in this situation, no further corrections for multiple comparisons were applied. The obtained t-values were thresholded by a t-value corresponding to a two-tailed significance level of 1% ⁵³.

Results

Demographics, pre-and post-training group differences in clinical scores

Of the ten patients enrolled in NFB training, one patient dropped out after six NFB sessions and another patient dropped out after 14 sessions (out of a total of 16 sessions). Table 1 summarizes the clinical findings. There were no significant between-group or within-group differences in the pre-post comparison of psychopathological symptoms (Table 2, I).

The P200 control group showed a tendency for symptom improvement with regard to PANSS score (pre-test vs. post-test, $t_{(4)} = 2.71$, $p = 0.054$). None of the cognitive test scores (HVLT-R, NAB-Mazes, or D2 test of attention) were significantly associated with any learning parameter.

Table 1

Pre- post differences in hallucinations

There were no consistent changes in the HCS on an overall level (Table 2, II): Independent of group, three subjects showed improvement, three showed no change, and three showed deterioration (Figure 3).

Figure 3

Group effects in pre- post differences in hallucinations

As shown in Figure 3, there were no significant differences in AVH symptom improvement (HCS difference score) between the treatment and control group ($t_{(7)} = -0.53$, $p = 0.62$) (Table 2, III).

Parametrization of individual NFB training patterns

Similarly to the already published analysis of the results obtained when the same training protocol was applied to healthy subjects (but not yet known when we conducted the training of the patients), a linear mixed model including a regressor for the first vs second session per day (Table 2, IV) indicated that training was ineffective in

patients in the second session ($\beta = -0.08$, $t = -1.89$, $p = 0.06$). As in the paper on the healthy subjects, the data from the second session per day was therefore not included in the remaining analyses.

The individual parameters of the linear mixed model explaining the NFB learning behavior, as well as the statistics on the β -values are shown in Figure 4. A significant negative effect of carry-over was identified. Decreases in amplitude of the trained ERP component were interpreted to be the result of a habituation effect^{54, 55}. Patients showed a trend of within-session training learning; however, this effect was non-significant.

The loading of the different NFB learning parameter on the extracted principal factor is shown in Table 3. Closer inspection of the load structure indicates that in our data, subjects primarily varied along a dimension showing, in parallel, more within-session training aptitude, within-session carry-over (amplitude increases in the passive condition) as well as across-session learning (amplitude increases in training relative to baseline), or the reverse. We will therefore refer to positive values of this load as sustained learning, whereas negative values will be referred to as transient learning.

Group comparisons of NFB training patterns

There were no significant effects of group on EEG-NFB performance across sessions (Group effect: $t = -1.22$, $p = 0.25$; Aptitude x Group effect: $t = -0.29$, $p = 0.78$; Day x Group effect: $t = 0.37$, $p = 0.72$; Learning x Group: $t = -0.00$, $p = 0.99$) or within sessions (Group effect: $t = -0.63$, $p = 0.54$; Aptitude x Group effect: $t = -1.12$, $p = 0.26$; Repetition x Group effect: $t = -0.83$, $p = 0.41$; Learning x Group effect: $t = 0.95$, $p = 0.35$). Based on this result, subsequent neurophysiological analyses were performed across all patients (i.e., independent of group) (Table 2, V).

Relationship of hallucination changes and NFB learning type

To test for a systematic relationship between NFB learning and AVH severity (Table 2, VI), we used the Pearson product-moment correlation value between the individual scores of the NFB learning factor and the HCS difference score across both groups (treatment and control). Results showed a significant effect ($r = 0.66$, $n = 9$, $p = 0.05$, Figure 5). Patients showing more of the sustained NFB learning pattern showed an improvement in the hallucinations compared to those patients that showed more of the transient learning pattern.

Breaking this further down, a separate correlation analysis for each group yielded a strong significant correlation between the learning factor and the change in hallucinations in the treatment group ($r = 0.97$, $n = 4$, $p = 0.03$), whereas no significant correlation was found in the control group ($p = 0.23$, $n = 5$, $p = 0.72$). However, a cautious interpretation is necessary due to the small group numbers. These results showed that a positive load in the learning factor was positively associated with an improvement in the hallucinations.

Figure 5 about here

Extraction of psychopathological characteristics

The analysis of the relationship of NFB learning and psychopathology was based on a factor analysis of the various psychopathology scores. The loading of these different scores on this psychopathology factor is shown in Table 3. The factor was primarily indicative of a negative content of AVH, a high amount of distress and absence of wellbeing, or the reverse (Table 2, VII).

Relationship of pre-training psychopathology and NFB learning type

To test whether pre-training symptom severity could predict NFB learning, factor scores obtained from clinical data were correlated with the scores of the NFB learning factor. Pearson product-moment correlation coefficients were non-significant ($r = 0.01$, $n = 10$, $p = 0.98$) (Table 2, VIII).

Pre-post differences in ERP topography and source estimates

A microstate analysis identified the exact time interval of the N100 component as 79–119 ms with typical central negativity, and that of the P200 component as 136–204 ms (Table 2, IX). For the N100, there was no significant spatial difference ($p = 0.28$) or GFP effect ($p = 0.14$) identified in the pre-post comparison. For the P200, there was a significant spatial pre-post difference ($p = 0.04$) and a tendency for higher GFP during post-training ($p = 0.07$). In the P200 interval, voxel-wise paired t-tests of sLORETA estimates were significant in the right PAC (Brodmann area [BA] 41 and 42); superior temporal gyrus (BA 22); inferior temporal gyrus, fusiform gyrus, and parahippocampal

gyrus (BA 20); posterior cingulate cortex (BA 29 and 30); parahippocampal gyrus (BA 35 and 36); and insular cortex (BA 13) (Figure 6).

Investigating group specific influences of the N100 training (treatment group) on the P200 component showed a significant topographical ($p = 0.002$) and GFP effect ($p = 0.002$) in the pre-post comparison, but not within the N100 component. However, P200 training (control group) did not significantly affect the N100 nor the P200 components in these analyses.

Figure 6 about here

Discussion

The goal of this study was to evaluate the utility of N100 EEG-NFB training for treating AVH symptoms in patients with treatment-resistant schizophrenia. Contrary to our hypothesis, NFB training did not afford significant improvements in schizophrenia psychopathology. The P200 control group showed a tendency for a reduction in the total score of the Positive and Negative Syndrome Scale (PANSS) after the last NFB training session compared to the first session. A possible explanation could be an effect of treatment expectation. This, however, does not explain an increase in the total PANSS score seen in the treatment group. Considering this, the tendency for a reduction in the total score of the PANSS in the control group might result from the fact that this group started with a higher score compared to the treatment group in the first place, and thus may be explained as a regression to the mean. Comparing the total score after the training in both group showed that, despite a tendency for a reduction, the control group still showed a higher score compared to the treatment group.

We did not identify relationships between individual psychopathological characteristics and NFB performance, indicating that symptom severity was not a predictor for NFB learning. This is in line with Gruzelier et al.²¹ who concluded that the psychotic symptoms of schizophrenia are not as obstructive for learning as other factors such as anxiety or loss of motivation in NFB tasks. Our result is consistent with this observation.

The use of a mixed effects model analysis allowed us to model NFB data with different learning parameters. With regard to NFB performance, Gruzelier et al.²¹ previously described the differentiation of within-session and across-

session learning. Diaz Hernandez et al. ³⁶ further distinguished between training aptitude and training learning within and across sessions. Assessing a learning factor based on the different learning parameters showed a dimensional distribution of NFB performance in the patients with opposing characteristics: More *sustained learning* was associated with more within-session aptitude and across-session learning, with less within-session learning and across-session aptitude and carry-over. More *transient learning* was related to an opposite pattern, exhibiting better within-session learning and across-session carry-over and aptitude. To further explore the significance of these learning parameters, we examined the relationship between these learning dynamics and AVH severity. As described above, comparing group averages in the psychopathology did not show a significant effect. A more accurate representation of psychopathological symptoms might however be seen in a dimensional approach ⁵⁶. Following this assumption, correlation measures were used to assess the relationship between NFB learning and AVH. There was a significant correlation between learning and change in hallucination severity in pre- and posttraining comparison, indicating that subjects rather showing *sustained learning* experienced a decrease in the perception of AVHs post-training. Within *sustained learning*, training aptitude was negatively associated with AVH severity. Training aptitude represented the general ability of a subject to increase the trained ERP in the training condition compared to the passive condition, independent of time (number of sessions). Therefore, within-session aptitude is not a measure of rapid learning, but is rather a quality that was inherent or developed slowly across repeated sessions. It is possible that subjects showing more *sustained learning* had a pre-existing advantage (within-session aptitude) that allowed them to improve their control of AVHs. Hence, within session aptitude may be an important predictor of NFB-training performance and therapeutic benefit. In summary of these findings, the target component of EEG-NFB (i.e., N100 or P200) did not appear to have any significance for the effect of NFB on AVH symptoms or control; rather, within-session aptitude was more relevant to therapeutic success. This observation can be explained by the hypothesis that a common mechanism was affected by the training of both components.

Finally, it is noteworthy that we identified a significant spatial difference in the source of P200 in our pre-training versus post-training comparison, mainly in temporal regions (i.e., the PAC and superior and inferior temporal gyri) and parts of the limbic system (i.e., posterior cingulate cortex and parahippocampal gyrus). The role of P200 is not yet fully understood ^{57, 58}; however, attention seems to influence its amplitude. Crowley and Colrain

⁵⁷ reported that a higher attention load led to a decrease in P200 amplitude. Tremblay et al. ⁵⁹ observed increases in P200 amplitude after auditory evoked potential exposure and concluded that this effect arose from task exposure or execution. Since we identified a spatial P200 effect that was independent of group (treatment or control) and NFB performance, it is assumed that P200 does not represent a biomarker for learning ⁵⁹.

The present study had some limitations. First, despite a modification of the P200 amplitude across training sessions, our data did not support an effect of EEG-NFB training on the N100 amplitude using our protocol. This may have been due to the fact that we tried to modify a component that was already disturbed in patients with schizophrenia. Alternatively, it is possible that we did not use a protocol that sufficiently enforced modification of the N100 amplitude (e.g. using a different beep tone interval, or omitting the need for strategy development).

Second, another limitation is given by the number of sessions needed in a clinical disorder. It has been previously shown that schizophrenia patients needed more training sessions than healthy subjects in order to achieve the same control in neurofeedback²². In addition, our data suggested that the second session of each day was not helpful for the training. Therefore, we suggest that future studies consider distributed learning across more training sessions for patient groups and at the same time avoiding too long training sessions or double sessions that outstrip the patients' capacity to concentrate and properly train with NFB.

Third, and relevant to our first limitation, we were only able to recruit a small number of participants despite a substantial two-year effort. Thus, a lack of statistical power may have driven our inability to identify effects of EEG-NFB and between-group differences. Taken together, we were unable to conclude whether targeting the N100 amplitude is a useful feature of EEG-NFB training. However, significant effects in the treatment group (N100 training) on the ERP P200 component indicated a change through the N100 NFB training, which was not found in the control group (P200 training). At the same time, these results do not interfere with the primary goal of affecting auditory verbal hallucinations through modulation of the N100. Nonetheless, the interpretation of this finding remains open and needs further investigation.

Future studies should adopt a NFB protocol that better targets the N100 amplitude, or alternatively consider the use of another method such as TMS or tDCS to modulate N100 amplitude. Moreover, further studies should explore the mechanism by which subjects with sustained learning experienced improvements in AVHs in our paradigm.

References

1. Curcic-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.
2. Dierks T, Linden DE, Jandl M, et al. Activation of Heschl's gyrus during auditory hallucinations. *Neuron.* 1999; 22: 615-21.
3. Dyck MS, Mathiak KA, Bergert S, et al. Targeting treatment-resistant auditory verbal hallucinations in schizophrenia with fMRI-based neurofeedback - exploring different cases of schizophrenia. *Front Psychiatry.* 2016; 7: 37.
4. Allen P, Modinos G, Hubl D, et al. Neuroimaging auditory hallucinations in schizophrenia: from neuroanatomy to neurochemistry and beyond. *Schizophr Bull.* 2012; 38: 695-703.
5. Tracy DK and Shergill SS. Mechanisms Underlying Auditory Hallucinations-Understanding Perception without Stimulus. *Brain Sci.* 2013; 3: 642-69.
6. Shergill SS, Murray RM and McGuire PK. Auditory hallucinations: a review of psychological treatments. *Schizophr Res.* 1998; 32: 137-50.
7. Lincoln T. *Kognitive Verhaltenstherapie in der Schizophrenie - Ein individuenzentrierter Ansatz.* 2. ed. Göttingen: Hogrefe, 2014.
8. Brunelin J, Mondino M, Gassab L, et al. Examining transcranial direct-current stimulation (tDCS) as a treatment for hallucinations in schizophrenia. *Am J Psychiatry.* 2012; 169: 719-24.
9. Homan P, Kindler J, Federspiel A, et al. Muting the voice: a case of arterial spin labeling-monitored transcranial direct current stimulation treatment of auditory verbal hallucinations. *Am J Psychiatry.* 2011; 168: 853-4.
10. Kindler J, Homan P, Flury R, Strik W, Dierks T and Hubl D. Theta burst transcranial magnetic stimulation for the treatment of auditory verbal hallucinations: results of a randomized controlled study. *Psychiatry Res.* 2013; 209: 114-7.
11. Sommer IE, Slotema CW, Daskalakis ZJ, Derks EM, Blom JD and van der Gaag M. The treatment of hallucinations in schizophrenia spectrum disorders. *Schizophr Bull.* 2012; 38: 704-14.
12. Strehl U. What learning theories can teach us in designing neurofeedback treatments. *Front Hum Neurosci.* 2014; 8: 894.
13. Arns M, Heinrich H and Strehl U. Evaluation of neurofeedback in ADHD: the long and winding road. *Biol Psychol.* 2014; 95: 108-15.
14. Gevensleben H, Rothenberger A, Moll GH and Heinrich H. Neurofeedback in children with ADHD: validation and challenges. *Expert Rev Neurother.* 2012; 12: 447-60.
15. Strehl U, Leins U, Goth G, Klinger C, Hinterberger T and Birbaumer N. Self-regulation of slow cortical potentials: a new treatment for children with attention-deficit/hyperactivity disorder. *Pediatrics.* 2006; 118: e1530-40.
16. Wangler S, Gevensleben H, Albrecht B, et al. Neurofeedback in children with ADHD: specific event-related potential findings of a randomized controlled trial. *Clin Neurophysiol.* 2011; 122: 942-50.
17. Schoenberg PL and David AS. Biofeedback for psychiatric disorders: a systematic review. *Appl Psychophysiol Biofeedback.* 2014; 39: 109-35.
18. Serman MB and Egnor T. Foundation and practice of neurofeedback for the treatment of epilepsy. *Appl Psychophysiol Biofeedback.* 2006; 31: 21-35.
19. Tan G, Thornby J, Hammond DC, et al. Meta-analysis of EEG biofeedback in treating epilepsy. *Clin EEG Neurosci.* 2009; 40: 173-9.
20. Thompson L, Thompson M and Reid A. Neurofeedback outcomes in clients with Asperger's syndrome. *Appl Psychophysiol Biofeedback.* 2010; 35: 63-81.
21. Gruzelier J, Hardman E, Wild J and Zaman R. Learned control of slow potential interhemispheric asymmetry in schizophrenia. *Int J Psychophysiol.* 1999; 34: 341-8.
22. Schneider F, Rockstroh B, Heimann H, et al. Self-regulation of slow cortical potentials in psychiatric patients: schizophrenia. *Biofeedback Self Regul.* 1992; 17: 277-92.

23. McCarthy-Jones S. Taking back the brain: could neurofeedback training be effective for relieving distressing auditory verbal hallucinations in patients with schizophrenia? *Schizophr Bull.* 2012; 38: 678-82.
24. Cordes JS, Mathiak KA, Dyck M, et al. Cognitive and neural strategies during control of the anterior cingulate cortex by fMRI neurofeedback in patients with schizophrenia. *Front Behav Neurosci.* 2015; 9: 169.
25. Hubl D, Koenig T, Strik WK, Garcia LM and Dierks T. Competition for neuronal resources: how hallucinations make themselves heard. *Br J Psychiatry.* 2007; 190: 57-62.
26. Rosburg T, Boutros NN and Ford JM. Reduced auditory evoked potential component N100 in schizophrenia--a critical review. *Psychiatry Res.* 2008; 161: 259-74.
27. Ford JM, Roach BJ, Palzes VA and Mathalon DH. Using concurrent EEG and fMRI to probe the state of the brain in schizophrenia. *Neuroimage Clin.* 2016; 12: 429-41.
28. Mulert C, Gallinat J, Pascual-Marqui R, et al. Reduced event-related current density in the anterior cingulate cortex in schizophrenia. *Neuroimage.* 2001; 13: 589-600.
29. Mulert C, Seifert C, Leicht G, et al. Single-trial coupling of EEG and fMRI reveals the involvement of early anterior cingulate cortex activation in effortful decision making. *Neuroimage.* 2008; 42: 158-68.
30. Edgar JC, Hanlon FM, Huang MX, et al. Superior temporal gyrus spectral abnormalities in schizophrenia. *Psychophysiology.* 2008; 45: 812-24.
31. Ford JM, Dierks T, Fisher DJ, et al. Neurophysiological studies of auditory verbal hallucinations. *Schizophr Bull.* 2012; 38: 715-23.
32. Wu KY, Chao CW, Hung CI, Chen WH, Chen YT and Liang SF. Functional abnormalities in the cortical processing of sound complexity and musical consonance in schizophrenia: evidence from an evoked potential study. *BMC Psychiatry.* 2013; 13: 158.
33. Busse M, Low YF, Corona-Strauss FI, Delb W and Strauss DJ. Neurofeedback by neural correlates of auditory selective attention as possible application for tinnitus therapies. *Conf Proc IEEE Eng Med Biol Soc.* 2008; 2008: 5136-9.
34. Delb W, Strauss DJ, Low YF, et al. Alterations in Event Related Potentials (ERP) associated with tinnitus distress and attention. *Appl Psychophysiol Biofeedback.* 2008; 33: 211-21.
35. Haller S, Birbaumer N and Veit R. Real-time fMRI feedback training may improve chronic tinnitus. *Eur Radiol.* 2010; 20: 696-703.
36. Diaz Hernandez L, Rieger K, Baenninger A, Brandeis D and Koenig T. Towards using microstate-neurofeedback for the treatment of psychotic symptoms in schizophrenia. A feasibility study in healthy participants. *Brain Topogr.* 2016; 29: 308-21.
37. Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia.* 1971; 9: 97-113.
38. Kay SR, Fiszbein A and Opler LA. The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. *Schizophrenia Bulletin.* 1987; 13: 261-76.
39. Haddock G, McCarron J, Tarrier N and Faragher EB. Scales to measure dimensions of hallucinations and delusions: the psychotic symptom rating scales (PSYRATS). *Psychol Med.* 1999; 29: 879-89.
40. Kern RS, Nuechterlein KH, Green MF, et al. The MATRICS Consensus Cognitive Battery, part 2: co-norming and standardization. *Am J Psychiatry.* 2008; 165: 214-20.
41. Nuechterlein KH, Green MF, Kern RS, et al. The MATRICS Consensus Cognitive Battery, part 1: test selection, reliability, and validity. *Am J Psychiatry.* 2008; 165: 203-13.
42. Brickenkamp R. *Test d2 - Aufmerksamkeits-Belastung-Test.* Göttingen: Germany: Hogrefe, 1981.
43. Hoffman RE, Gueorguieva R, Hawkins KA, et al. Temporoparietal transcranial magnetic stimulation for auditory hallucinations: safety, efficacy and moderators in a fifty patient sample. *Biol Psychiatry.* 2005; 58: 97-104.
44. Rieger K, Rarra MH, Moor N, et al. Neurofeedback-Based Enhancement of Single Trial Auditory Evoked Potentials: Feasibility in Healthy Subjects. *Clin EEG Neurosci.* 2017: 1550059417708935.
45. Peirce JW. PsychoPy--Psychophysics software in Python. *J Neurosci Methods.* 2007; 162: 8-13.
46. Team RC. R: A language and environment for statistical computing. *R Foundation for Statistical Computing.* Vienna, Austria 2013.

47. Addington D, Addington J and Maticka-Tyndale E. Specificity of the Calgary Depression Scale for schizophrenics. *Schizophr Res.* 1994; 11: 239-44.
48. Evans C, Connell J, Barkham M, et al. Towards a standardised brief outcome measure: psychometric properties and utility of the CORE-OM. *Br J Psychiatry.* 2002; 180: 51-60.
49. Murray MM, Brunet D and Michel CM. Topographic ERP analyses: a step-by-step tutorial review. *Brain Topogr.* 2008; 20: 249-64.
50. Strik WK, Fallgatter AJ, Brandeis D and Pascual-Marqui RD. Three-dimensional tomography of event-related potentials during response inhibition: evidence for phasic frontal lobe activation. *Electroencephalogr Clin Neurophysiol.* 1998; 108: 406-13.
51. Kikuchi M, Koenig T, Wada Y, et al. Native EEG and treatment effects in neuroleptic-naive schizophrenic patients: time and frequency domain approaches. *Schizophr Res.* 2007; 97: 163-72.
52. Pascual-Marqui RD. Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. *Methods Find Exp Clin Pharmacol.* 2002; 24 Suppl D: 5-12.
53. Hubl D, Schneider RC, Kottlow M, et al. Agency and ownership are independent components of 'sensing the self' in the auditory-verbal domain. *Brain Topogr.* 2014; 27: 672-82.
54. Naatanen R and Picton T. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology.* 1987; 24: 375-425.
55. Rosburg T and Soros P. The response decrease of auditory evoked potentials by repeated stimulation--Is there evidence for an interplay between habituation and sensitization? *Clin Neurophysiol.* 2016; 127: 397-408.
56. Strik W, Wopfner A, Horn H, et al. The Bern psychopathology scale for the assessment of system-specific psychotic symptoms. *Neuropsychobiology.* 2010; 61: 197-209.
57. Crowley KE and Colrain IM. A review of the evidence for P2 being an independent component process: age, sleep and modality. *Clin Neurophysiol.* 2004; 115: 732-44.
58. Woodman GF. A brief introduction to the use of event-related potentials in studies of perception and attention. *Atten Percept Psychophys.* 2010; 72: 2031-46.
59. Tremblay KL, Ross B, Inoue K, McClannahan K and Collet G. Is the auditory evoked P2 response a biomarker of learning? *Front Syst Neurosci.* 2014; 8: 28.

Tables:

Table 1: Demographic and psychopathological data

	Treatment (N100)		Control (P200)		Statistics	
	n = 4		n = 6		Treatment vs. Control	
	M	SD	M	SD	coeff	p-value
Demographics						
age (years)	37,2	18,2	36,2	13,5	t = -0.11	p = 0.92
education (years)	10,3	1,5	11,7	2,1	t = 1.17	p = 0.28
medication (CE)	333,3	245,0	616,5	389,1	t = 1.28	p = 0.24
Psychopathology						
PsyRats pre	29,5	9,3	32,0	6,9	t = 0.47	p = 0.66
PsyRats post	29,0	7,4	31,8	6,4	U = 9.0	p = 0.61
PANSS pre	49,5	8,1	65,0	13,4	t = 1.75	p = 0.12
PANSS post	56,3	13,2	58,8	13,7	t = 0.28	p = 0.78
	coeff	p-value	coeff	p-value		
PsyRats: pre - post	Z = 0.0	p = 1.0	Z = -0.55	p = 0.58		
PANSS: pre - post	t = -2	p = 0.14	t = 2.71	p = 0.054		

CE = Chlorpromazine equivalents [mg], M = mean; SD = standard deviation; coeff = coefficient; PsyRats = Psychotic Rating Scale; PANSS = Positive And Negative Syndrome Scale

Table 2: Overview of the performed analyses

	Pre-and post training state		Pre-post difference		IV: Individual NFB training patterns (parametrization using linear mixed models and PCA)	
	<i>Group-wise</i>	<i>Group-wise</i>	<i>Overall</i>	<i>Group-wise</i>	<i>Overall</i>	
						V: Group comparisons of NFB training patterns (linear mixed models)
Demographics and clinical scores	I: Demographics, pre- and post-training group differences in clinical scores (t-tests / U-tests)					
Change in hallucinations (assessed by the HCS)		III: Group effects in pre- post differences in hallucinations (t-tests)	II: Pre- post differences in hallucinations (t-tests)			VI: Relationship of hallucination changes and NFB learning type
VII: Extraction of psychopathological characteristics (PCA)						VIII: Relationship of pre-training psychopathology and NFB learning type
Evoked potentials			IX: Pre-post differences in ERP topography and source estimates.			

Table 3: Factor extraction for neurophysiological and clinical measures

Neurophysiological measures	Factor
WS: training-aptitude: constant difference between N100/P200 in the training compared to passive condition averaged within	0,901
AS: training-learning: increase of the N100/P200 in the training compared to the passive condition across session	0,888
WS: carry-over: non-voluntary change of the N100/P200 amplitude in the passive condition within sessions	0,841
AS: transfer-aptitude: constant difference between N100/P200 in the transfer compared to passive condition averaged across session	0,521
AS: intercept: between subject offset across session	0,383
WS: intercept: between subject offset within session	-0,179
AS: training-aptitude: constant difference between N100/P200 in the training compared to passive condition averaged across	-0,584
AS: transfer-learning: increase of the N100/P200 in the training compared to the passive condition across session	-0,720
AS: carry-over: non-voluntary change of the N100/P200 amplitude in the passive condition across sessions	-0,768
WS: training-learning: increase of the N100/P200 in the training compared to the passive condition within session	-0,860
Clinical measures	Factor
PR: negative content (amount)	0,898
PR: distress (amount)	0,891
PR: control (frequency)	0,886
PR: frequency	0,842
PR: loudness	0,811
PR: distress (intensity)	0,779
PR: disruption	0,729
PR: duration	0,709
PR: negative content (degree)	0,637
PANSS pre negative	0,632
PR: control (influence)	0,451
PR: location	0,351
PANSS pre positive	0,102
PR: beliefs re-origin	-0,008
CDSS: pre totalscore	-0,087
wellbeing	-0,233

AS = Across Session; WS = Within Session; PR = PsyRats; CDSS = Calgary Depression Scale Schizophrenia

Figures:

Figure 1. Overview of the experimental procedure

Figure 2: Schematic overview of the pre-session and NFB training, Abbreviations: AEP, auditory evoked potential.

Figure 3. Hallucination Change Scale scores at the first training session (HCS pre) and thirteenth training session (HCS post)

Figure 4. Overview of individual subject performance for each learning parameter within and across sessions

Figure 5. Correlation between NFB learning and change in hallucinations

Figure 6. Significant pre-training versus post-training spatial differences ($p < 0.01$) in the P200 component