


# SCIENTIFIC REPORTS



OPEN

## Power-spectra and cross-frequency coupling changes in visual and Audio-visual acquired equivalence learning

András Puszta<sup>1</sup> , Ákos Pertich<sup>1</sup>, Xénia Katona<sup>1</sup>, Balázs Bodosi<sup>1</sup>, Diána Nyujtó<sup>1</sup>, Zsófia Giricz<sup>1</sup>, Gabriella Eördegh<sup>2</sup> & Attila Nagy<sup>1</sup>

The three phases of the applied acquired equivalence learning test, i.e. acquisition, retrieval and generalization, investigate the capabilities of humans in associative learning, working memory load and rule-transfer, respectively. Earlier findings denoted the role of different subcortical structures and cortical regions in the visual test. However, there is a lack of information about how multimodal cues modify the EEG-patterns during acquired equivalence learning. To test this we have recorded EEG from 18 healthy volunteers and analyzed the power spectra and the strength of cross-frequency coupling, comparing a unimodal visual-guided and a bimodal, audio-visual-guided paradigm. We found that the changes in the power of the different frequency band oscillations were more critical during the visual paradigm and they showed less synchronized activation compared to the audio-visual paradigm. These findings indicate that multimodal cues require less prominent, but more synchronized cortical contribution, which might be a possible biomarker of forming multimodal associations.

Associative learning is a complex task in which rule-transfer is increased between two superficially dissimilar stimuli (or antecedents) that have previously been associated with similar outcomes (or consequents)<sup>1</sup>. Catherine E. Myers and co-workers developed a learning paradigm (Rutgers Acquired Equivalence Test, also known as the fish-face paradigm), which can be applied to investigate a specific kind of associative learning, visually-guided equivalence learning<sup>2</sup>. This complex test consists of three phases which can be interpreted as better-known tests: The initial phase can be described as an associative learning or trial-and-error learning or rule-based learning task, which primarily requires an intact basal ganglia-network<sup>2,3</sup>. The second part is the retrieval phase, which can be interpreted as working memory maintenance, and the third part is the generalization or rule-transfer phase, which primarily requires an intact hippocampal-mediotemporal system<sup>4,5</sup>. Our research group adapted and modified the Rutgers Acquired Equivalence Test to make it more sensitive, and recently investigated the development of these learning functions in healthy humans and how migraines affect them<sup>6</sup>. This learning paradigm critically requires the normal function of subcortical structures, i.e. hippocampi and basal ganglia<sup>2,7</sup>. Cortical contribution is also necessary in the visually-guided learning paradigm as stimulus representations and associations are stored and (re)activated in stimulus-relevant cortical areas<sup>8–11</sup>. Thus, associative learning requires cooperation between the learning circuit and other task-specific brain areas.

Different areas of the brain must interact to provide the basis for the integration of sensory information, sensory-motor coordination and many other functions that are critical for learning, memory, and perception. Hebb suggested that this is accomplished by forming assemblies of cells whose synaptic linkages are strengthened whenever the cells are activated synchronously<sup>12</sup>. Neuronal oscillations are a natural consequence of forming such cell assemblies via the summation of hundreds of EPSPs and IPSPs, and the cerebral cortex generates multitudes of oscillations at different frequencies mainly through inhibiting spike-trains at a specific frequency. To investigate these oscillations during associative learning, our research group used electroencephalography (EEG),

<sup>1</sup>Department of Physiology, Faculty of Medicine, University of Szeged, Dóm tér 10, Szeged, Hungary. <sup>2</sup> Department of Oral Biology and Experimental Dental Research, Faculty of Dentistry, University of Szeged, Tisza Lajos krt. 64, Szeged, Hungary. Correspondence and requests for materials should be addressed to A.P. (email: [puszta.andris@gmail.com](mailto:puszta.andris@gmail.com)) or A.N. (email: [nagy.attila.1@med.u-szeged.hu](mailto:nagy.attila.1@med.u-szeged.hu))

which is a well-known non-invasive monitoring method for investigating the electrical signals generated by large assemblies of neurons and their connections.

A number of investigations have described the EEG-features of the different phases of associative learning and memory. One key feature of the reward-related learning is that the positive feedback elicits beta power increment, while negative feedback causes power increment in both theta and beta power<sup>13,14</sup>. Furthermore, studies in associative learning tasks revealed gamma coherence over parietooccipital areas<sup>15,16</sup>. In working memory tasks, frontal midline theta power increment is a well-known phenomenon<sup>17–19</sup>. Theta/alpha-gamma cross frequency coupling in working memory load was described earlier<sup>20,21</sup>. The tasks described above mainly used unimodal stimulus-pairs to associate, and there is less information about how these patterns change if we apply multimodal stimulus-pairs.

It is well known from earlier studies that both brain structures fundamentally involved in visual associative learning, the basal ganglia and the hippocampi, receive not only visual but also multisensory information<sup>22–25</sup>. A bimodal or multimodal stimulus could be more informative in its complexity than a unimodal stimulus from the environment. The studies referenced above investigated visually-guided equivalence learning and to our knowledge no study has addressed the cortical contribution to multisensory-guided acquired equivalence learning.

Multimodal information could be more informative, than a unimodal stimulus from the environment<sup>26,27</sup>. Multisensory integration occurs at different levels of brain functions. It can be observed at the cellular level<sup>28–31</sup> in several brain regions, such as the superior colliculus<sup>32</sup>, the basal ganglia<sup>33,34</sup>, the cortex<sup>35</sup>, and the hippocampus<sup>36</sup>, and it can also be observed on the behavioral level<sup>37,38</sup>. In the present study we investigate the effect of multisensory stimuli on associative learning and we ask whether there are any specific changes in the power spectra and in the cross-frequency coupling of the human neocortex in a multisensory task compared to a unimodal visual one<sup>39</sup>.

## Results

Altogether 23 healthy volunteers participated in the investigation. For the biomathematical analysis (including the psychophysical results, time-frequency (TF) results, cross-frequency coupling results, and power - /synchronization index (SI) - performance correlation), the raw electrophysiological data of 18 volunteers were analyzed, as in the other recordings the signal to noise ratio was low, and neither the excessive attempt to clean the data from muscular and ocular artefacts with preprocessing methods described earlier could make them acceptable.

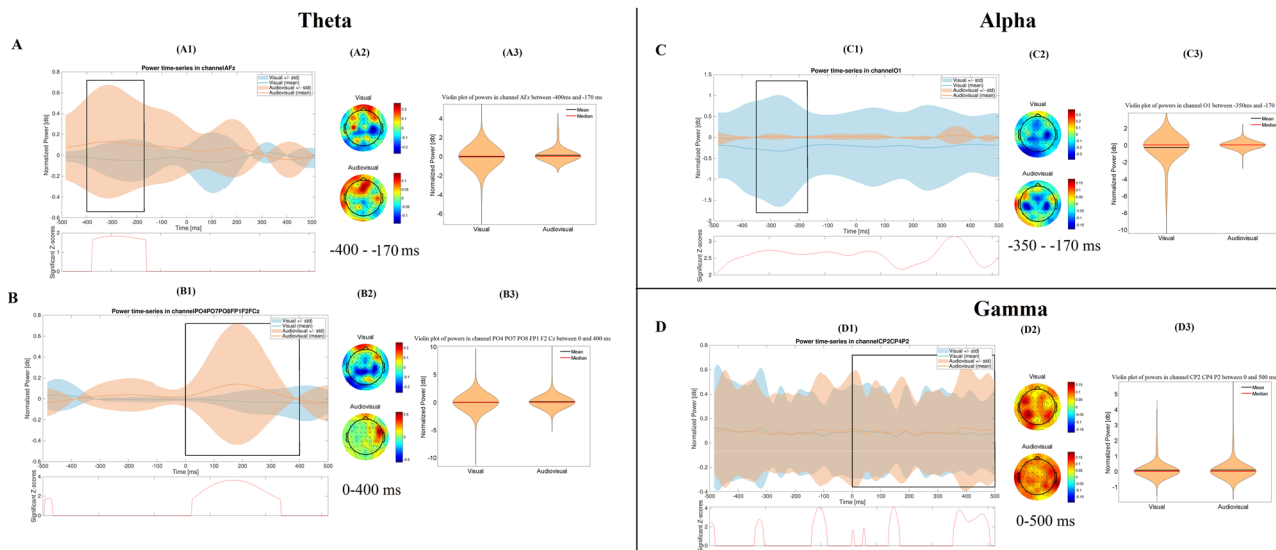
**Data visualization.** The electrophysiological results in four different frequency bands (theta (4–7 Hz), alpha (8–13 Hz), beta 14–30 Hz), and gamma (31–70 Hz) will be presented below for each phase (Acquisition, Retrieval and Generalization) of the two (visual and Audio-visual) paradigms.

In the time-frequency results, the group-level statistical differences between the time-frequency power spectra of the visual and audiovisual paradigm are presented in each frequency band, and in each phase of the paradigm (See Figs 1–3). In every case, the results are discussed in a range –500 ms–500 ms in case of the Acquisition phase, and –500 ms–0 ms in case of the Retrieval and Generalization phases, where the 0 ms denotes the time of the answer. We will give a detailed description of the statistical differences between the visual and the audiovisual paradigm only in those cases, where we found significant difference between the visual and the audiovisual paradigms. Detailed descriptions of the changes in the time-frequency power spectra in each phase of the visual and Audio-visual paradigm compared to baseline activity are presented in the Supplementary Data (Supplementary Data 1). Furthermore, because of the huge amount of data, our detailed results of the time-frequency analysis cannot be interpreted with one plot. Instead, we provide an interactive surface, where the significant changes are available on each of the 64 channels in 10 ms-time bins (Supplementary Data 1).

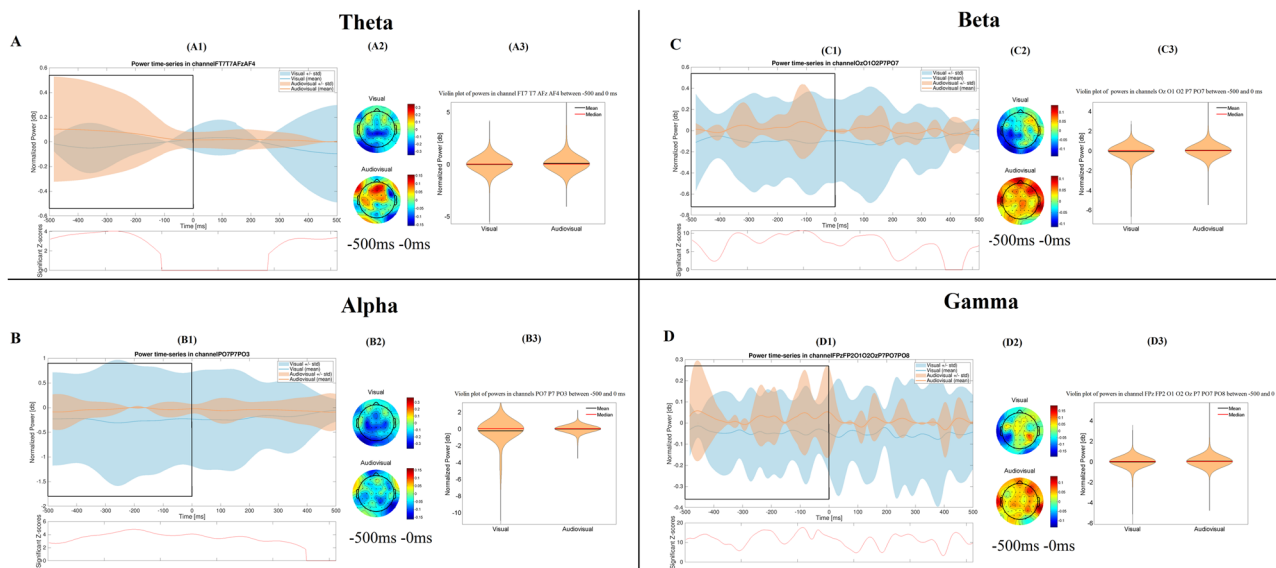
In the cross-frequency coupling results, the group-level statistical differences between the mean synchronization indices of the visual and Audio-visual paradigm are presented in each frequency band, and in each phase of the paradigm. Also, detailed descriptions of the SI-value changes of each phase of the visual and Audio-visual paradigm compared to the SI-value of the baseline activity are presented. Significant changes of the cross-frequency coupling in each channel in different phases of the paradigm will be presented using the *topoplot* function of EEGLab. For plotting purposes, only significant changes (i.e. where the Z-scores were  $> 1.69$ ) are presented on the plots (Fig. 4). In the smaller topographical plots, the significant difference between SI-values of the given phase and the background activity is shown. The red color indicates where the SI was significantly higher during the given phase compared to baseline-activity, and the blue color indicates where the SI in the given phase was significantly lower compared to baseline activity. On the larger topographical plots, we present the significant difference between the visual and audiovisual paradigm. Here, the red color indicates that the power of that specific frequency band in the given phase of the paradigm was significantly higher during the audiovisual task compared to the visual task, where the blue color indicates the opposite. As we found that the highest changes in the individual comodulogram occurred at the modulating frequency band 8–15 Hz and the modulated frequency band at 31–45 Hz, our results will indicate the group level results found in that frequency range.

In the power-performance-correlation results, significant correlations (i.e. where the t-values were  $> 2.583$ ) are presented in each phase of the visual and audiovisual paradigm. For plotting purposes, only significant correlations (i.e. Z-score  $> 1.69$ ) between the performance in the psychophysical test and the cortical power changes (Fig. 5) are plotted on the topographical figures.

**Performance in the visual and the Audio-visual psychophysical tests.** The mean correct trial ratios (correct trials/all trials) during different phases of the visual acquired equivalence test were as follows: 0.92 in the acquisition phase (range = 0.83–0.98, SD  $\pm$  0.04), 0.98 in the retrieval phase (range = 0.9–1, SD  $\pm$  0.02), and 0.98 in the generalization phase (range = 0.92–1, SD  $\pm$  0.04), (Fig. 6A). Repeated measure analysis of variance (ANOVA) revealed significant difference in the correct trial ratios ( $F = 20.87$ ,  $p < 0.001$ ), and Tukey post-hoc



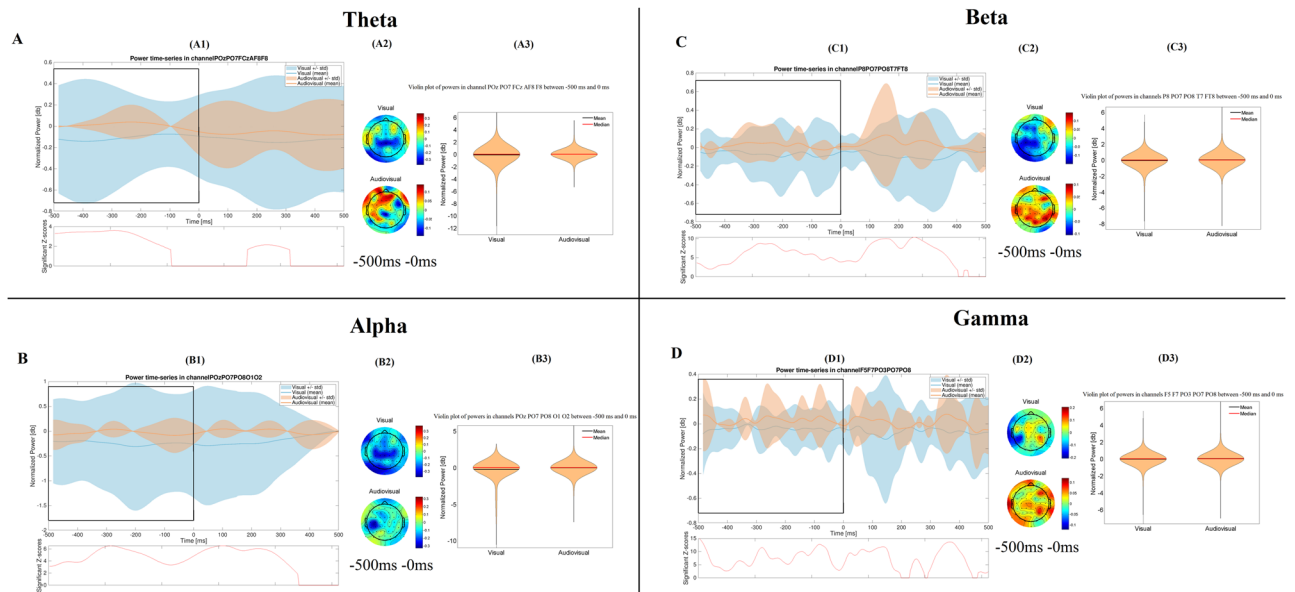
**Figure 1.** Time-frequency results in the acquisition phase. The figure shows the most important differences, which were found between the visual and the audiovisual tasks. Part (A) and Part (B) represent the results in the theta band, before and after the given answer, respectively. The (C,D) parts show the results in the alpha and gamma bands, respectively. Within each part of the figure, there are three different subplots. The subplots with number 1 (A1,B1,C1,D1) show the normalized power-fluctuation of the given channels, and the significant Z-scores between the two time-series calculated with random permutation test before and after 500 ms of the given answer. 0 ms denotes the time point when the answer was given. The subplots with number 2 (A2,B2,C2,D2) show the topographical representation of the mean normalized power in certain time-windows. The red colour indicates power-increase, while the blue one indicates power-decrease compared to baseline-activity. The subplots with 3 (A3,B3,C3,D3) show the violin plot of the normalized powers in the selected time-window and in the selected channels.



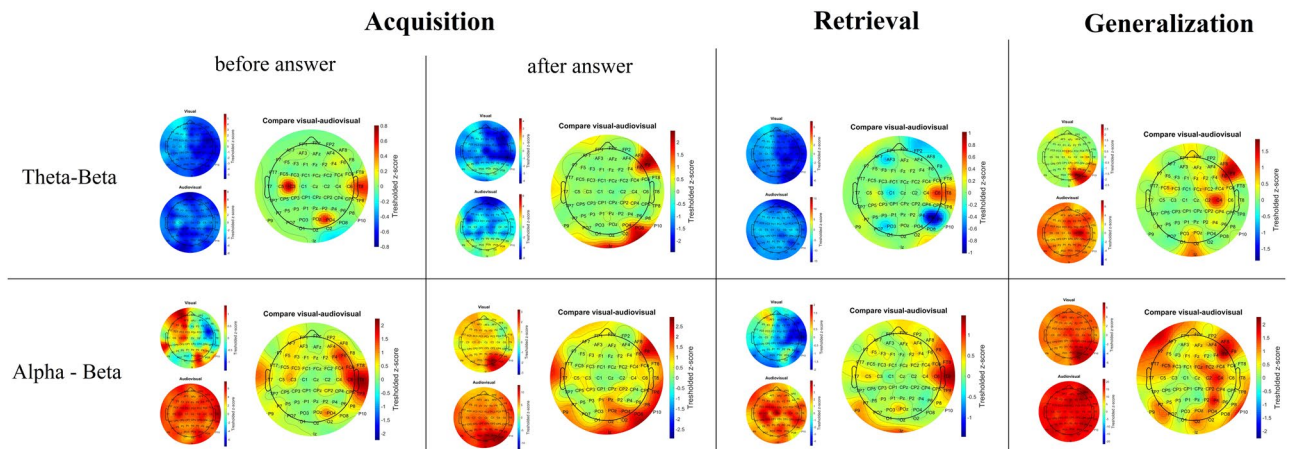
**Figure 2.** Time-frequency results in the retrieval phase. The figure shows the most important differences, which were found between the visual and the audiovisual task. The (A–D) parts show the results in the theta, alpha, beta and gamma bands, respectively. Other conventions are same as on Fig. 1.

analysis revealed that the correct trial ratio in the acquisition phase was significantly lower compared to the retrieval and generalization phases ( $p < 0.001$ ). The correct trial ratios did not differ significantly between the generalization and retrieval phase ( $p = 0.992$ ).

The mean correct trial ratios during different phases of the Audio-visual acquired equivalence test were as follows: 0.94 in the acquisition phase (range = 0.90–0.98,  $SD \pm 0.02$ ), 0.98 in the retrieval phase (range = 0.88–1,



**Figure 3.** Time-frequency results in the generalization phase. The figure shows the most important differences, which were found between the visual and the audiovisual task. The (A–D) parts show the results in the theta, alpha, beta and gamma bands, respectively. Other conventions are same as on Fig. 1.

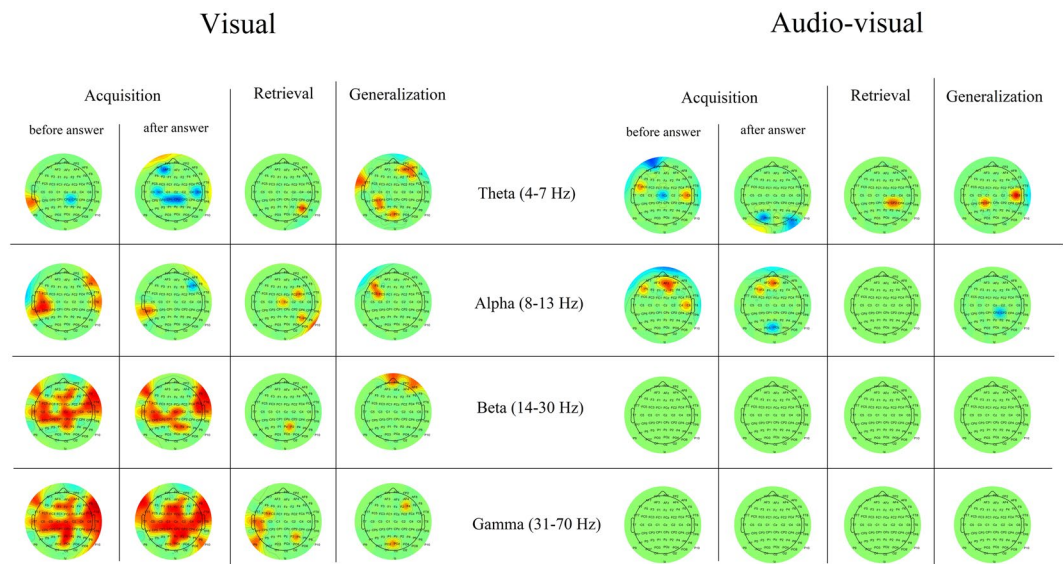


**Figure 4.** Topographical representation of the cross-frequency theta-beta and alpha-beta coupling results during the visual and audiovisual task. In the smaller topographical plots, the significant difference between SI-values of the given phase and the background activity is shown. The red color indicates where the SI was significantly higher during the given phase compared to baseline-activity, and the blue color indicates where the SI in the given phase was significantly lower compared to baseline activity. On the larger topographical plots, we present the significant difference between the visual and audiovisual paradigm. Here, the red color indicates that the power of that specific frequency band in the given phase of the paradigm was significantly higher during the audiovisual task compared to the visual task, where the blue color indicates the opposite.

SD ± 0.03), and 0.97 in the generalization phase (range = 0.83–1, SD ± 0.05), (Fig. 6B). Repeated measure ANOVA revealed significant difference in the correct trial ratios ( $F = 7.49, p = 0.002$ ), and Tukey post-hoc analysis revealed that the correct trial ratio in the acquisition phase was significantly lower compared to the retrieval ( $p = 0.002$ ) and generalization phases ( $p = 0.019$ ). The correct trial ratios did not differ significantly between the generalization and retrieval phase ( $p = 0.709$ ).

**Group-level analysis of the time-frequency power spectra in the visual and Audio-visual associative learning paradigms.** We interpret our EEG results in each frequency band by showing the cortical areas and the corresponding channels where we found significant differences between the power spectra of the two paradigms. Because of the huge amount of data, our detailed results cannot be interpreted with one plot. Instead, we provide an interactive surface where each significant difference is presented in all 64 channels and in





**Figure 5.** Topographical representation of the power-performance correlation in each phase of the paradigm, during the visual and audiovisual task. Significant correlations (i.e. where the  $t$ -values were  $>2.583$ ) are presented in each phase of the visual and audiovisual paradigm. For plotting purposes, only significant correlations between the performance in the psychophysical test and the cortical power changes are plotted on the topographical figures.

each 10 ms-time section (Supplementary Data 1). Additionally, we provide summarizing pictures and a table of our results (see in Figs 1–3 for Acquisition, Retrieval, Generalization, respectively and Table 1).

**Acquisition phase.** For the graphical interpretation of our time-frequency results in the acquisition phase see Fig. 1.

The Mann-Whitney test revealed that the power of the theta band was significantly higher during the audiovisual paradigm (mean = 0.118 dB, STD = 0.5 dB, Range = 0 dB 2.814 dB) compared to the visual paradigm (mean = -0.042 dB, STD = 0.459 dB, Range = -3.698 dB 2.012 dB) over the frontal channels, 400 ms to 170 ms before the answer. After the answer (from 0 ms to 400 ms after the answer), the power of the theta band was significantly higher ( $p < 0.001$ ) in the audiovisual paradigm (mean = 0.078 dB, STD = 0.494 dB, Range = -1.943 dB 6.794 dB) compared to the visual paradigm (mean = -0.023 dB, STD = 0.655 dB, Range = -6.954 dB 5.212 dB) not only over the frontal but over the parietooccipital channels, too.

In case of the alpha frequency band we found that the power was significantly lower ( $p < 0.001$ ) during the visual paradigm (mean = -0.278 dB, STD = 1.159 dB, Range = -6.664 dB 0 dB), than in the audiovisual one (mean = 0.012 dB, STD = 0.140 dB, Range = -1.386 dB 1.041 dB) over the occipital channels, 350 ms to 170 ms before the answer.

We observed no significant difference in the beta power between the visual and the audiovisual paradigm.

The power of the gamma band was significantly higher ( $p = 0.005$ ) during the audiovisual paradigm (mean = 0.093 dB, STD = 0.417 dB, Range = -0.496 dB 4.472 dB) than in the visual one (mean = 0.08 dB, STD = 0.36 dB, Range = -0.842 dB 3.491 dB), over the parietal channels, starting from 0 ms until 500 ms after the given answer.

**Retrieval phase.** For the graphical interpretation of our time-frequency results in the retrieval phase see Fig. 2.

The Mann-Whitney test revealed that the power of the theta band was significantly higher ( $p < 0.001$ ) in the audiovisual paradigm (mean = 0.066 dB, STD = 0.383 dB, Range = -1.871 dB 3.718 dB) than in the visual paradigm (mean = -0.02 dB, STD = 0.276 dB, Range = -3.549 dB 2.006 dB) over the temporal and frontal channels, 500 ms to 0 ms before the answer.

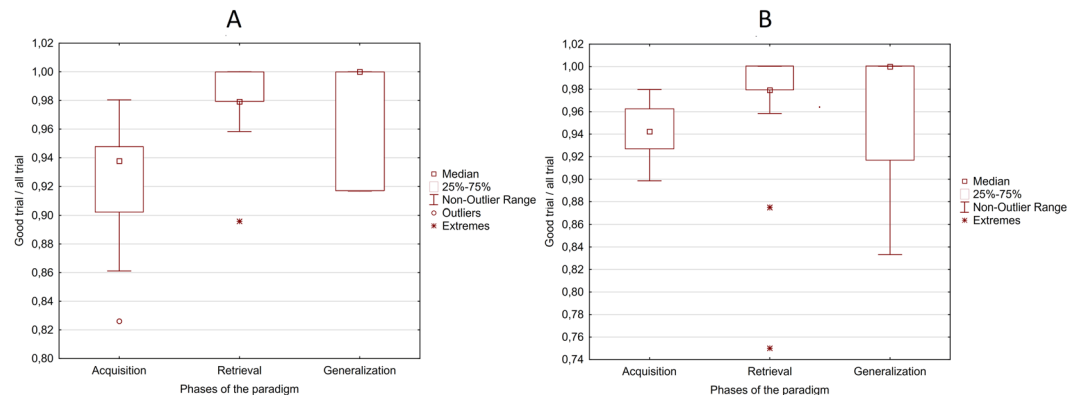
In case of the power of the alpha frequency band we found that it was significantly lower ( $p < 0.001$ ) during the visual paradigm (mean = -0.253 dB, STD = 1.058 dB, Range = -8.308 dB 0 dB) than in the audiovisual paradigm (mean = -0.036 dB, STD = 0.222 dB, Range = -2.324 dB 0.971 dB) over the parietooccipital channels, from 500 ms before the answer.

The power of the beta frequency band was significantly higher ( $p < 0.001$ ) during the audiovisual paradigm (mean = 0.027 dB, STD = 0.276 dB, Range = -3.61 dB 2.994 dB) than in the visual paradigm (mean = -0.093 dB, STD = 0.47 dB, Range = -5.524 dB 1.11 dB), over the occipital and parietooccipital channels, from 500 ms before the answer.

The power of the gamma band was significantly higher ( $p < 0.001$ ) during the audiovisual paradigm (mean = 0.026 dB, STD = 0.27 dB, Range = -3.205 dB 4.133 dB) compared to the visual paradigm

Condition	Frequency band	Time bounds	Cortical region	Power in visual paradigm				p-value	Power in audiovisual paradigm			
				mean	std	min	max		mean	std	min	max
Acquisition	Theta	-400ms -170ms	Frontal	-0,042	0,459	-3,698	2,012	<0.001	0,118	0,500	0,000	2,814
Acquisition	Theta	0 ms 400 ms	Frontal Parietooccipital	-0,023	0,655	-6,954	5,212	<0.001	0,078	0,494	-1,943	6,794
Acquisition	Alpha	-350ms -170ms	Occipital	-0,278	1,159	-6,644	0,000	<0.001	0,012	0,140	-1,386	1,041
Acquisition	Gamma	0 ms 500 ms	Parietal	0,080	0,360	-0,842	3,491	0.005	0,093	0,417	-0,496	4,472
Retrieval	Theta	-500 ms 0 ms	Temporal Frontal	-0,020	0,276	-3,549	2,006	<0.001	0,066	0,383	-1,871	3,718
Retrieval	Alpha	-500 ms 0 ms	Parietooccipital	-0,253	1,058	-8,308	0,000	<0.001	-0,036	0,222	-2,324	0,971
Retrieval	Beta	-500 ms 0 ms	Occipital Parietooccipital	-0,093	0,470	-5,524	1,110	<0.001	0,027	0,276	-3,610	2,994
Retrieval	Gamma	-500 ms 0 ms	Frontal + Parietooccipital	-0,041	0,302	-4,662	2,088	<0.001	0,026	0,270	-3,205	4,133
Generalization	Theta	-500 ms 0 ms	Frontal + Parietooccipital	-0,107	0,677	-8,810	2,529	<0.001	0,009	0,356	-3,022	3,157
Generalization	Alpha	-500 ms 0 ms	Occipital + Parietooccipital	-0,245	1,065	-8,110	0,381	<0.001	-0,033	0,409	-4,785	3,102
Generalization	Beta	-500 ms 0 ms	Parietooccipital	-0,066	0,438	-6,089	3,068	<0.001	0,017	0,347	-5,557	3,933
Generalization	Gamma	-500 ms 0 ms	Frontal + Parietooccipital	-0,038	0,376	-5,521	3,490	<0.001	0,026	0,346	-4,803	4,537

**Table 1.** Table of the statistics of the most important differences between the visual and audiovisual time-frequency results. We provide a summarizing table of the descriptive statistic of the normalized powers in different frequency band and time-window identified in the interactive surface provided in the Supplementary Data 1. The p-value between the visual and audiovisual task was calculated with Mann-Whitney test.



**Figure 6.** Box plots of the correct trial ratios in each phase of the paradigm during visual (A) and Audio-visual (B) acquired equivalence learning test.

(mean = -0.041 dB, STD = 0.302 dB, Range = -4.662 dB 2.088 dB), over the frontal and parietooccipital channels, starting from 500 ms before the answer.

**Generalization phase.** For the graphical interpretation of our time-frequency results in the generalization phase see Fig. 3.

The Mann-Whitney test revealed that the power of the theta band was significantly higher ( $p < 0.001$ ) during the audiovisual paradigm (mean = 0.009 dB, STD = 0.356 dB, Range = -3.022 dB 3.157 dB) than in the visual one (mean = -0.127 dB, STD = 0.677 dB, Range = -8.810 dB 2.529 dB) over the frontal and parietooccipital channels, from 500 ms before the answer.

The power of the alpha frequency band was significantly lower ( $p < 0.001$ ) during the visual paradigm (mean = -0.245 dB, STD = 1.065 dB, Range = -8.11 dB 0.381 dB), compared to the audiovisual paradigm (mean = -0.033 dB, STD = 0.409 dB, Range = -4.785 dB 3.102 dB) over the occipital and parietooccipital channels, from 500 ms before the answer.

The power of the beta frequency band was significantly higher ( $p < 0.001$ ) during the audiovisual paradigm (mean = 0.017 dB, STD = 0.347 dB, Range = -5.557 dB 3.102 dB) compared to the visual paradigm (mean = -0.066 dB, STD = 0.438 dB, Range = -6.089 dB 3.068 dB), over the parietooccipital channels, starting from 500 ms before the answer.

The power of the gamma band was significantly higher in the audiovisual paradigm (mean = 0.026 dB, STD = 0.346 dB, Range = -4.803 dB 4.537 dB) than in the visual paradigm (mean = -0.038 dB, STD = 0.376 dB, Range = -5.521 dB 3.49 dB), over the frontal and parietooccipital channels, starting from 500 ms before the answer.

**Correlation between TF-power and psychophysical performance.** To reveal the significant TF-power that correlated with the psychophysical performance, we calculated the correlation between the mean power of different frequency bands and performance in the different phases of the paradigm (Acquisition – before and after the given answer – Retrieval, Generalization). See detailed description of the calculations in the Materials and Methods section. In this section we summarize the significant differences in the regions in different frequency bands during the different phases of the paradigm. In the topographical figures we present the significant correlation coefficients found in different phases of the paradigm (Fig. 5).

**Visual associative learning paradigm.** We found a positive correlation between the power of the higher frequency oscillations (beta, gamma) during the acquisition phase over the parietal and temporal channels before and after the answer.

In the alpha band, we found positive correlations between the power over the temporal channels before and after the answer and performance.

In the theta band power, we found a positive correlation between the performance and power before the answer over the temporal channels. However, we found that the power of theta band over the parietal channels after the answer was negatively correlated with performance.

During the retrieval phase of the visual-associative learning paradigm, we found that power of the gamma band over the temporal channels positively correlated with performance. Concerning the lower frequency bands, their power over the parietooccipital channels positively correlated with performance.

In case of the generalization phase, we observed positive correlation between the psychophysical performance and the power of beta band over the frontal channels.

In case of the alpha band, we observed positive correlation over the frontal channels, and in case of the theta band we found positive correlation over the frontotemporal and parietal channels.

**Audiovisual associative learning paradigm.** In the acquisition phase of the audiovisual paradigm we found no significant correlation between the power of the higher frequency band (beta, gamma) and the psychophysical performance.

We observed positive correlation between the power of the alpha band before the answer over the frontal and frontotemporal channels and performance. After the answer we found, that the power of the alpha band positively correlated over the frontal channels, and negatively correlated over the parietooccipital channels.

We observed negative correlation over the frontal channels and positive correlation over the temporoparietal channels between performance and the power of the theta band before the answer. After the answer we found, that the power of the theta band negatively correlated over the parietooccipital channels with performance.

In the retrieval phase of the audiovisual paradigm we found, that the power of the theta band over the parietal channels positively correlated with the psychophysical performance. We did not find any significant correlations between the power of other frequency bands and the psychophysical performance.

In the generalization phase we found, that the power of the alpha band over the parietooccipital negatively correlated with performance. We also found that the power of the theta band over the parietotemporal channels positively correlated with performance.

### **Cross-frequency coupling in the visual and Audio-visual associative learning paradigms.**

During the acquisition phase of the paradigm, we found a significant decrease in cross-frequency theta-beta coupling compared to baseline-activity both in the case of the visual and the audiovisual paradigm before the given answer. After the given answer, we found significant SI-elevation compared to baseline activity over the occipital-parietooccipital channels (Iz, PO8), which was higher during the audiovisual paradigm.

Regarding the alpha-beta coupling, we found a significant increase of the SI-values before the answer over the whole scalp in the case of the audiovisual paradigm, and over the left frontal-frontotemporal (FP1, AF3, AF7, F1, F3, F5, F7, FC1, FC3, FC5) and parietooccipital-occipital (PO4, O2) channels. Comparing the visual and the audiovisual paradigm, before the given answer in the acquisition phase, we found that the SI-values were significantly higher over the temporal (T7, T8, TP8, C6) and frontotemporal (FC6, F6) channels during the audiovisual paradigm.

After the answer, we found that alpha-beta coupling was significantly higher compared to the baseline activity both in the visual and the audiovisual paradigm, predominantly over the parietooccipital-occipital (PO4, PO8, O2, Oz) channels, which was higher during the audiovisual paradigm.

During the retrieval phase of the paradigm, we observed significant theta-beta coupling decrease over the whole scalp during both the visual and audiovisual paradigm. Comparing the visual and the audiovisual paradigm, we found that the SI-decrease was higher over the right parietooccipital (P8, P6) channels during the visual paradigm.

For the alpha-beta coupling, we observed a significant decrease of the SI-values compared to baseline activity over the right temporal, frontotemporal and parietooccipital channels during the visual paradigm. We observed significant alpha-beta cross-frequency coupling increase over the whole scalp during the audiovisual paradigm. Comparing the visual and the audiovisual paradigm, we found that the SI-values were higher over the right temporal-frontotemporal (T8, C6, FT8, F8) and parietooccipital-occipital (Iz, P9) channels during the audiovisual paradigm.

During the generalization phase, we found significant theta-beta coupling over the parietooccipital-occipital channels (PO8, PO4, P2, P4, P6, P8) during the visual paradigm. We found that the theta-beta coupling was significantly higher compared to baseline activity over the whole scalp during the audiovisual paradigm. Comparing

the visual and audiovisual theta-beta coupling strength, we found that the SI-values were significantly higher over the frontotemporal (C4, F8, F6) and occipital (Iz, Oz, POz) channels during the audiovisual paradigm.

In the case of the alpha-beta coupling, we found the same pattern of SI-value increase that was observed in the theta-beta coupling during the generalization phase.

## Discussion

The present study analyzed the EEG correlates in a visually-guided and an Audio-visually (bimodal or multisensory) guided acquired -equivalence learning tasks. To our knowledge, this is the first study to address the comparison of the cortical power spectra and their changes in a unimodal visual and a multisensory associative learning task. The major finding of the study is that the cortical activity depends critically on the phase of the paradigm, and some changes in cortical powers are characteristic to unimodal visual and multisensory Audio-visual tasks. In general, during the audiovisual paradigm, the power changes of the event-related low and high-oscillations were higher compared to the visual paradigm, but the psychophysical performance of the acquisition phase only correlated with the power of different frequency bands during the visual paradigm. On the other hand, while the power changes of the event-related oscillations were higher during the audiovisual paradigm, the performance did not depend on the power of different oscillations, and the strength of the cross-frequency coupling was higher. Furthermore, the performance of the acquisition phase seems to be more connected to the strength of the alpha-beta coupling during the audiovisual paradigm. We are convinced that the cortical power differences in the two paradigms cannot be the result of having previously completed the first task (precondition), hence the order of the two paradigms (visual and Audio-visual) varied randomly across subjects.

The role of multimodal cues in associative learning has been widely investigated (for review see<sup>40</sup>). The psychological studies provided evidence, that multisensory working memory improves recall for cross-modal objects compared to modality-specific objects<sup>41,42</sup>, working memory capacity is higher for cross-modal objects under certain circumstances<sup>43</sup> and visual and auditory information can interfere with each other<sup>44</sup>. While former studies revealed mainly cortical areas are involved in associative learning<sup>45,46</sup>, only few electrophysiological studies showed the functional basis of the multisensory integration<sup>47</sup>, and to our knowledge, our study is the first that describes the role of different oscillations in multisensory integration during learning.

The performance of the investigated population in the psychophysical test (acquisition error ratio, retrieval error ratio, generalization error ratio) was in the same range as that of the earlier investigated healthy controls of neurological and psychiatric patients<sup>2,4,6,48</sup>. Based on this we are strongly positive that the electrophysiological results showed here are representative.

One of the common findings both in the visual and the Audio-visual paradigm is the increased theta band activity in the parietooccipital, frontal midline, and prefrontal areas during the acquisition and retrieval phases. Frontal midline theta activity has been widely investigated (for review, see<sup>49</sup>), and its contribution seems to be obvious to internally-guided cognitive tasks that require no external responses<sup>50–52</sup>. A more general interpretation of the increased theta power in the frontal cortex could be the coordinated reactivation of information represented in visual areas. This was also found in single-unit recordings in primate V4<sup>53</sup> as well as LFP-synchronisation between the prefrontal cortex and V4<sup>54</sup>. Regarding our findings, we hypothesize that the initial acquisition phase of the task requires more repeated reactivation of the cortical areas where the stimulus is processed. We also assume (based on the power-performance correlation we observed) that the better the associations are encoded, the more enhanced theta activity can be observed in the frontal midline areas.

The earlier findings of human electrophysiological studies indicated the role of the alpha band in visual<sup>55,56</sup> as well as Audio-visual processing<sup>57</sup>. Moreover, Hanslmayr and his colleagues<sup>58</sup> found that the performance in processing stimuli is more likely to be linked to decreased power in the alpha band. Another function that has been attributed to alpha activity is a mechanism of sensory suppression, thus functional gating of information in the task-irrelevant brain areas<sup>59,60</sup>. Indeed, our findings suggest that the initial parts of the trials (0–50 ms) are coupled to decreased alpha power, which were then followed by an increase of it in the visual (occipital) and Audio-visual (parietooccipital) cortical areas. These together suggest that after a rapid processing of the cue image and sound, the threshold of the visual/Audio-visual cortical areas for external stimuli becomes higher, allowing the information to be encoded (or retrieved) internally.

In the case of the beta frequency range, there is a growing evidence that enhanced beta oscillations appear in patients with Parkinson disease (PD)<sup>61</sup>. A number of investigations found that a power increase of the beta frequency band correlates with the severity of parkinsonian motor symptoms such as akinesia and rigidity<sup>62,63</sup>. As a result, beta oscillations are currently investigated as a potential biomarker tracking the effectiveness of deep brain stimulation treatment of PD patients<sup>64</sup>. Furthermore, deep brain stimulation of the subthalamic nucleus at beta frequencies worsens motor symptoms in PD patients<sup>65</sup>. One of the main functions of the basal ganglia is to contribute to associative learning by the trial-and-error method<sup>66,67</sup>. It is well-known that in PD (along with other deficits in which basal ganglia are affected) this learning mechanism is reduced<sup>33,34</sup>. Having seen in our results that a robust decrease of beta-power occurred in all phases (acquisition, retrieval and generalization phases) of both the visual and Audio-visual learning tasks, one may consider that this cortical power density decrease in the beta band is a necessary cortical outcome of the normal action of the basal ganglia in visual and Audio-visual associative learning.

The gamma frequency band plays an important role in memory processes<sup>68</sup> as well as other cognitive processes, such as word learning, reading, and expectancy<sup>36,37</sup>. We observed a power increase during the acquisition phase of the task in the frontal cortex and in associative cortical areas connected to the modality of the presented stimuli (i.e. the occipital cortex in the visual task, the parietotemporal areas in the Audio-visual task). Thus, we hypothesize that the acquisition phase of the learning paradigm needs strong cortical contribution. The power differences between the acquisition phase of the visual and Audio-visual learning paradigm suggest stronger cortical contribution to the multisensory learning task. On the other hand, this increase in the gamma power was not



detectable in the retrieval and the generalization phases of the paradigm. The explanation for this could be that the already-learned acquisitions were already transmitted to the hippocampus and the application of the earlier acquisitions does not need strong cortical activation in the gamma band. There is also evidence that a decrease in the gamma power of the local field potential correlates with performance and attention by selectively gating sensory inputs<sup>69,70</sup>. We assume that the decrease of the gamma power we found during the task over the cortical areas where the stimulus was processed (i.e. occipital areas in the visual and parietotemporal in the Audio-visual task) could be beneficial in memory encoding and retrieval by filtering irrelevant external stimuli.

Calculating the synchronization indices, we found increased coupling between theta and alpha/beta in each phase of both the visual and the Audio-visual task, which are in accordance with earlier studies that emphasize the role of theta-gamma/beta coupling in memory processes<sup>40,41</sup> as well as the alpha gamma/beta coupling in visual perception<sup>42,43</sup>. We also found that this synchronization was significantly stronger during each phase of the Audio-visual task than in the visual paradigm. Furthermore, during the retrieval and generalization phases, we found that the synchronization between theta-beta and alpha-beta was significantly stronger in the audiovisual task. We argue that the audiovisual paradigm required stronger synchronized cortical contribution in all phases of the task because of the cross-modal integration of the visual and auditory stimuli. Furthermore, as the power of the cortical oscillations increased more during the visual task compared to the Audio-visual task, this indicates that the multimodal (audiovisual) associations require less, but more synchronized activation of the cortex.

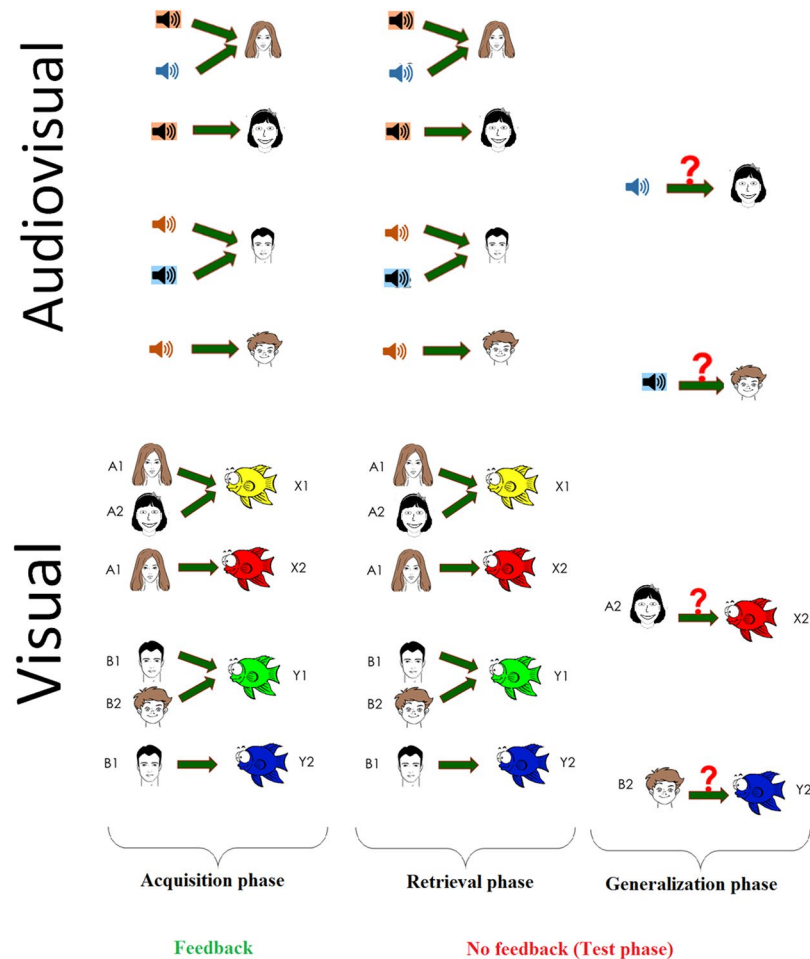
Optimal performance in the acquisition phase of the learning paradigms appears to depend mainly on the integrity of the basal ganglia, whereas performance in the test phase (both retrieval and generalization) has been linked to the integrity of the hippocampal region<sup>2,71,72</sup>. It is also known that both fundamentally-involved structures, the basal ganglia and the hippocampi, are involved in attention processes (for a reviews, see<sup>73</sup>). At the behavioral level, multisensory integration could be dependent on the level of attention and is not an automatic, unconscious process<sup>74</sup>. A multisensory task seems to be more complex and probably needs more attention from the participants. However, our psychophysical results show no significant differences between the performances in the visual and the multisensory tasks. In summary, although the role of attention cannot be excluded in the psychophysical learning test, the same level of performance in the unimodal visual and multimodal (Audio-visual) learning paradigms contradicts the assumption that attention contributes significantly to the differences in cortical activation patterns.

We can conclude that the changes in the power of the different frequency-band oscillations were more critical during the visual paradigm. On the other hand, the encoding and the retrieval part of the bimodal, Audio-visual task required more strongly synchronized cortical activity than those of the visual one. The two former statements are probably due to the fact that in the case of the multisensory associations the investigated memory processes (encoding, retrieval, and generalization) require less prominent, but more synchronized cortical activation, while the unimodal associations require more prominent and less synchronized cortical activity during the same memory processes. These findings further emphasize the effect of multimodal integration during associative learning and memory processes.

## Materials and Methods

The EEG data of 23 adult healthy young adults were recorded (12 females, 11 males, mean age: 26 years, range = 18–32). The participants were free of any ophthalmological or neurological conditions, and they were tested for parachromatism with the Pseudiso-chromatic Plate Color Vision Test. The participants were recruited on a voluntary basis. The potential subjects were informed about the background and goals of the study, as well as about the procedures involved. It was also emphasized that, given the lack of compensation or any direct benefit, the participants were free to quit at any time without any consequence (no one did so). Those who decided to volunteer signed an informed consent form. The study protocol conformed to the tenets of the Declaration of Helsinki in all respects, and was approved by the Medical Ethics Committee of the University of Szeged, Hungary (Number: 50/2015-SZTE). The datasets generated and analyzed during the present study and the Matlab codes, that were used in the analytical process connected to this study are available in the Supplementary Data 2.

**Visual associative learning test.** The testing software (described in earlier studies and originally written for iOS<sup>2</sup>) was adapted to Windows. It was coded in Assembly for Windows and translated into Hungarian, with the written permission of the copyright holder. The paradigm was also slightly modified to reduce the probability of completing its acquisition phase by mere guessing (see below). The tests were run on a PC. The stimuli were displayed on a standard 17-inch CRT monitor (refresh rate 100 Hz) in a quiet room separated by a one-way mirror from the recording room. Participants sat at a 114 cm distance from the monitor. One participant was tested at a time and no time limitation was set. The test was structured as follows: in each trial of the task, the participants saw a face and a pair of fish of different color, and had to learn through trial and error which fish was connected with which face (Fig. 1). There were four faces (A1, A2, B1, B2) and four possible fish (X1, X2, Y1, Y2), referred to as antecedents and consequents, respectively. In the initial, acquisition stages, the participants were expected to learn that when A1 or A2 appears, the correct answer was to choose fish X1 over fish Y1; given face B1 or B2, the correct answer was to choose fish Y1 over fish X1. If the associations were successfully learned, participants also learned that face A1 and A2 were equivalent with respect to the associated fish (faces B1 and B2 likewise). Next, participants learned a new set of pairs: given face A1, they had to choose fish X2 over Y2, and given face B1, fish Y2 over X2. This was the end of the acquisition phase. To this point, the computer provided feedback about the correctness of the choices, and six of the possible eight fish-face combinations were taught to the participants. In the following phases (retrieval and generalization), no feedback was provided. Beside the already-acquired six pairs (tested in the retrieval phase) the hitherto not shown two pairs were also presented, which were predictable based on the learned rules (tested in the generalization phase). Having learned, that faces A1 and A2 are equivalent, participants were expected to generalize from learning that if A1 goes with X2, A2 also goes with X2; the



**Figure 7.** Graphic overview of the unimodal visual (lower panel) and the bimodal Audio-visual (upper panel) acquired equivalence paradigms.

same holds for B2 (equivalent to B1) and Y2 (equivalent to B1). During the acquisition stages, new associations were introduced one by one, mixed with trials of previously-learned associations. The subjects had to achieve a certain number of consecutive correct answers after the presentation of each new association (4 after the presentation of the first association, and 4, 6, 8, 10, 12 with the introduction of each new association, respectively) to be allowed to proceed. In order to minimize the repetition effect in the acquisition phase, the last, 12-answer trial of the acquisition phase were set to be the part of the retrieval phase. This resulted in an elevated number of the required consecutive correct trials compared to the original paradigm, which made getting through the acquisition phase by mere guessing less probable. Similarly, in the test phase there were 48 trials (12 trials of new and 36 trials of previously-learned associations), as opposed to the 16 trials of the original paradigm.

**Audio-visual associative learning test.** We developed the Audio-visual (multisensory or bimodal) guided acquired equivalence learning test. The structure of the paradigm was the same as of the visual associative learning test, with the difference that the four antecedents were four sounds (A1, A2, B1, B2) and the consequents were the same four faces as in the visual associative learning paradigm. The main task of the participants was to determine from trial to trial which of the two given faces corresponds to the sound heard at the beginning of the trial. The sounds of the paradigm were a female voice saying “Hello”, the sound of a guitar, the sound of a motorcycle, and the sound of a cat. Each sound lasted less than 1 sec. The category rule implemented in the four faces (i.e. sex, age, hair color) was the same as in the visual paradigm, so similarity across the sounds was not important, but the similarity across the four faces was the same as in case of the visual paradigm. During the acquisition phase, six of the possible eight sound-face combinations were learned. During the test phase, no feedback was provided anymore, but beside the already-acquired six pairs (learned in the retrieval phase), the hitherto not shown last two pairs were also presented (generalization phase). For visual representation of the two task, see Fig. 7.

**Data acquisition.** Sixty-four channel EEG recordings were performed using a Biosemi ActiveTwo AD-box with 64 active electrodes. Actview software was used to set up the parameters and record the EEG-data (Biosemi B.V., The Netherlands). The sampling rate was 2048 Hz. Electrode offsets were all kept within normal and acceptable ranges. Raw signals were recorded on the computer that controlled the psychophysical learning task. The

stimulating software generated trigger signals to indicate the beginning of each trial. These trigger signals were recorded on an additional (sixty-fifth) channel. In order to obtain the baseline activity, one-minute-long resting state activities were recorded before and after testing the visual and Audio-visual associative learning test. The order of the two tests (visual and Audio-visual associative learning test) varied randomly across volunteers.

**Preprocessing.** The raw EEG data was first high-pass filtered (the filtering method used a 2 Hz highpass, two-way, least-squares FIR procedure, as implemented in the `eegfilt.m` script included in the `eeglab` package), and then re-referenced to the average of the channels. All trials were visually inspected, and those containing EMG or other artefacts not related to blinks were manually removed. Independent components analysis was computed using the EEGLab toolbox for Matlab<sup>75</sup>, and components containing blink/oculomotor artefacts or other artefacts that could be clearly distinguished from brain-driven EEG signals were subtracted from the data. The trials were defined as 500 ms before and after the given answer. In order to minimize the repetition effect in the acquisition phase, the last 12-answer trial of the acquisition phase was set to be the part of the retrieval phase. The subtracted trials did not exceed 1% of all trials. The mean trials of the phases of the paradigm was 54, 50 and 12, in the acquisition, retrieval and generalization phase, respectively. Additionally, noisy channels were interpolated using EEGLab toolbox. Then we used Laplacian fitting to improve the spatial resolution of the recording<sup>76</sup>. After the preprocessing steps, the data was resampled to 256 Hz, to minimize the computational time. This is the reason that in the present study the results of the time-frequency analysis are given in 10 ms-bins.

**Data analysis.** All data analysis was performed using Matlab (MATLAB and Statistics Toolbox Release 2018a, The MathWorks, Inc., Natick, Massachusetts, United States.) and Statistica software (Dell Statistica for Windows v13).

**Analysis of the performances in the psychophysical learning tasks.** The psychophysical data were analysed in three groups: data from the acquisition phase, data from the retrieval parts of the test phase (i.e. when the participant was presented an already-learned association), and data from the generalization part of the test phase (i.e. previously-not-learned associations). The number of correct and wrong responses were calculated in all phases, as well as the ratio of these to the total number of trials during the respective phase.

**Time-frequency analysis.** Time-frequency analysis was performed using a Continuous Morlet wavelet convolution (CMW) via FFT algorithm<sup>77</sup>. Firstly, FFT was first performed on one selected channel of the raw data. Then complex Morlet wavelets were created for each frequency (1–70 Hz) on which FFT was also executed. The cycles of the wavelets increased logarithmically as the frequency varied in a linear manner. After that, we calculated the dot product of the given channel's FFTs and the FFTs of the complex Morlet wavelets at each individual frequency, which yielded 70 complex numbers. Thereafter, the inverse Fast Fourier transform of the results of the dot product showed the alterations of the power in the time domain as follows:

$$K_x = IFFT(fft(C) \cdot fft(W_x))$$

where the  $K$  is the time-series of the given channel, wavelet-filtered to frequency  $x$ ,  $C$  is the time series of all trials of different phases, and  $W$  is the complex Morlet wavelet in a given frequency  $x$ . In order to avoid the edge-artifacts of the Morlet wavelet convolution, the raw data was multiplied five times before the convolution, yielding a two-series-long buffer zone at the beginning and the end of the time-series, which was cut out after the time-frequency analysis. After that, the channel's data was cut into different phases of the paradigm (baseline, acquisition, retrieval, generalization). The trials were defined as the 500 ms before and after the given answer. As the baseline activity was longer than the compared periods (i.e. the signal belonging to a given condition), the baseline activity was bootstrap-resampled to match the given condition by cutting one-second-long periods randomly from the baseline activity (1–1 minute before and after the first and the last trigger-signal, respectively). The bootstrapping method was the same as described in<sup>78</sup>. The data set for the purposed null-hypothesis (global band) was generated by iteratively calculating the mean difference of the randomized permutation of the power values of a particular channel in a given frequency band in two different phases of the paradigm (Baseline-Acquisition, Baseline-Retrieval, Baseline-Generalization). The Z-scores for each channel were then calculated between the distributions derived from the global band and the mean difference of the power values in a given frequency band between the analysed phase and the baseline activity. Z-scores were corrected by the minimum and maximum point of the null hypothesis distribution (also known as cluster-mass correction<sup>77,79</sup>).

Group-level analysis of the CMW was carried out in the same way as in the individual analysis described above, with the difference that the random permutation was performed across the mean power values of the subjects and not across the power value of each individual trial. The visualized methodological procedures of the above-described permutation-based test can be seen in one of our earlier publications through an individual example<sup>80</sup>. In the interactive surface provided in the Supplementary Material (Supplementary Data 1) only significant Z-scores calculated the above mentioned way are provided. As the data was resampled to 256 Hz, the time-bins in the Supplementary Data are 10 ms.

We identified the time-windows in which we found significant difference between the visual and the audio-visual paradigm, using the interactive surface provided in the Supplementary Material (Supplementary Data 1). After we identified the significant time-windows and the corresponding channels in each frequency band and condition, we additionally tested if the individual normalized powers of the different frequency bands in the selected channels and time-points are significantly different in the visual and the audiovisual paradigm by using Mann-Whitney test. The individual normalized powers were obtained by normalizing each individual time-frequency power in each condition and on each channel to the mean power of the baseline activity in the same channel and same frequency using decibel-normalization. Furthermore, to see the significant differences in

the time-domain in the selected channels, we performed permutation-test between the normalized power time series of the visual and audiovisual test.

**Calculation of event-related cross-frequency coupling.** Event-related synchronization index (SI) was calculated in order to examine whether the power of the high-frequency oscillations are coupled to the phase of the low-frequency oscillations on the same channel. The calculation method was almost the same as described by Cohen<sup>81</sup>. We will give a detailed description of the calculation of the SI in one phase of the paradigm in one channel's data referred as raw analytic signal. In the first step, the higher-frequency power time series were extracted from the concatenated trials. This was done by the combination of band-pass filtering and Hilbert transformation. First, we used a narrow band pass to filter the analytic signal to each frequency of beta and gamma band (15–70 Hz). The filtering method used a 4 Hz-width, two-way, least-squares FIR procedure (as implemented in the `eegfilt.m` script included in the `eeglab` package). Then we performed Hilbert transformation on the narrow bandpass-filtered epochs. The power time-series was extracted as the squared magnitude of  $z(t)$ , the analytic signal obtained from the Hilbert transform (power time series:  $p(t) = \text{real}[z(t)]^2 + \text{imag}[z(t)]^2$ ).

Then we band-pass filtered the raw analytic signal to each frequency of the low-frequency range (2–20 Hz, with 4 Hz-width). The phase of the band-pass filtered low- and high-frequency power time series were obtained from the Hilbert transform of the two time-series, respectively.

The synchronization between the phases of the two power time series can be calculated using the synchronization index (SI) as follows:

$$SI = \left| \frac{1}{n} \times \sum_{t=1}^n e^{i(\varphi_{ht} - \varphi_{lt})} \right|^2$$

where  $n$  is the number of time points,  $\varphi_{ht}$  is the phase value of the fluctuations in the higher-frequency power time series at time point  $t$ , and  $\varphi_{lt}$  is the phase value of the lower-frequency band time series at the same time point. The SI varied between 0 and 1. At SI 0 the phases are completely desynchronized, and at SI 1 the phases are perfectly synchronized.

Significant changes of the cross-frequency coupling at a population level were calculated by comparing the mean synchronization index in a given modulating - and modulated frequency range of the baseline activity and the given phase of the paradigm. The mean SI-values in each phase of the paradigm were then compared using permutation-based statistics, and the resulting Z-scores were corrected by the minimum and maximum point of the null hypothesis distribution (also known as cluster-mass correction<sup>77,79</sup>). See Supplementary Figs 1 and 2 for the graphical interpretation of the calculation of the SI values and their statistical comparison.

### Correlation between performance in the psychophysical test and the power density changes.

Correlation between individual performance and power density changes in a given channel and frequency band was also calculated in each phase of the paradigm. Performance was defined as the ratio of the correct (good) trials to all trials, and the individual power changes were defined as the individual Z-scores between the baseline activity's power density and the given phase's power density in a given channel in a given frequency band. The Pearson correlation coefficient was calculated using the `'corr'` function of Matlab. Statistical analysis was performed by calculating the t-score for each correlation coefficient as follows:

$$t_{ch,fr} = r_{ch,fr} * \sqrt{\frac{n-2}{1-r_{ch,fr}^2}}$$

where  $r$  is the correlation coefficient in channel ( $ch$ ) and frequency ( $fr$ ) band,  $n$  is the number of samples (in this case it was 18), and  $t$  is the calculated t value in a given channel and frequency band. T-values whose absolute values were smaller than 2.583 (which is the critical t-value if the degree of freedom is 16 and the significance level is 0.01) were set to 0. The corrected t-values in different frequency bands in each phase of the paradigm were plotted to a topographical map using the `'topoplot'` function of EEGLab.

### References

- Hall, G., Ray, E. & Bonardi, C. Acquired equivalence between cues trained with a common antecedent. *Journal of Experimental Psychology: Animal Behavior Processes* **19**, 391 (1993).
- Myers, C. E. *et al.* Dissociating hippocampal versus basal ganglia contributions to learning and transfer. *Journal of Cognitive Neuroscience* **15**, 185–193 (2003).
- Vadhan, N. P. *et al.* Stimulus–response learning in long-term cocaine users: Acquired equivalence and probabilistic category learning. *Drug & Alcohol Dependence* **93**, 155–162 (2008).
- Bódi, N., Csibri, E., Myers, C. E., Gluck, M. A. & Kéri, S. Associative learning, acquired equivalence, and flexible generalization of knowledge in mild Alzheimer disease. *Cognitive and Behavioral Neurology* **22**, 89–94 (2009).
- Kéri, S., Nagy, O., Kelemen, O., Myers, C. E. & Gluck, M. A. Dissociation between medial temporal lobe and basal ganglia memory systems in schizophrenia. *Schizophrenia research* **77**, 321–328 (2005).
- Óze, A. *et al.* Acquired equivalence and related memory processes in migraine without aura. *Cephalgia* **37**, 532–540 (2017).
- Shohamy, D. & Wagner, A. D. Integrating memories in the human brain: hippocampal-midbrain encoding of overlapping events. *Neuron* **60**, 378–389 (2008).
- Nyberg, L., Habib, R., McIntosh, A. R. & Tulving, E. Reactivation of encoding-related brain activity during memory retrieval. *Proceedings of the National Academy of Sciences* **97**, 11120–11124 (2000).
- Nyberg, L. *et al.* Reactivation of motor brain areas during explicit memory for actions. *Neuroimage* **14**, 521–528 (2001).
- Rösler, F., Heil, M. & Hennighausen, E. Distinct cortical activation patterns during long-term memory retrieval of verbal, spatial, and color information. *Journal of Cognitive Neuroscience* **7**, 51–65 (1995).
- Slotnick, S. D. Memory for color reactivates color processing region. *NeuroReport* **20**, 1568–1571 (2009).



12. Hebb, D. O. The Organization of Behavior: A Neuropsychological Theory (1949).
13. Donamayor, N., Marco-Pallarés, J., Heldmann, M., Schoenfeld, M. A. & Münte, T. F. Temporal dynamics of reward processing revealed by magnetoencephalography. *Human brain mapping* **32**, 2228–2240 (2011).
14. Cohen, M. X., Elger, C. E. & Ranganath, C. Reward expectation modulates feedback-related negativity and EEG spectra. *Neuroimage* **35**, 968–978 (2007).
15. Gruber, T., Keil, A. & Müller, M. M. Modulation of induced gamma band responses and phase synchrony in a paired associate learning task in the human EEG. *Neuroscience letters* **316**, 29–32 (2001).
16. Gruber, T., Müller, M. M. & Keil, A. Modulation of induced gamma band responses in a perceptual learning task in the human EEG. *Journal of cognitive neuroscience* **14**, 732–744 (2002).
17. Hsieh, L.-T. & Ranganath, C. Frontal midline theta oscillations during working memory maintenance and episodic encoding and retrieval. *Neuroimage* **85**, 721–729 (2014).
18. Jensen, O. & Tesche, C. D. Frontal theta activity in humans increases with memory load in a working memory task. *European journal of Neuroscience* **15**, 1395–1399 (2002).
19. Onton, J., Delorme, A. & Makeig, S. Frontal midline EEG dynamics during working memory. *Neuroimage* **27**, 341–356 (2005).
20. Schack, B., Vath, N., Petsche, H., Geissler, H.-G. & Möller, E. Phase-coupling of theta–gamma EEG rhythms during short-term memory processing. *International Journal of Psychophysiology* **44**, 143–163 (2002).
21. Roux, F. & Uhlhaas, P. J. Working memory and neural oscillations: alpha–gamma versus theta–gamma codes for distinct WM information? *Trends in cognitive sciences* **18**, 16–25 (2014).
22. Bates, S. L. & Wolbers, T. How cognitive aging affects multisensory integration of navigational cues. *Neurobiology of aging* **35**, 2761–2769 (2014).
23. Nagy, A., Eördegh, G., Paróczy, Z., Márkus, Z. & Benedek, G. Multisensory integration in the basal ganglia. *European Journal of Neuroscience* **24**, 917–924 (2006).
24. Ravassard, P. *et al.* Multisensory control of hippocampal spatiotemporal selectivity. *Science* **340**, 1342–1346 (2013).
25. Nagy, A., Paróczy, Z., Norita, M. & Benedek, G. Multisensory responses and receptive field properties of neurons in the substantia nigra and in the caudate nucleus. *European Journal of Neuroscience* **22**, 419–424 (2005).
26. Sarter, N. B. J. I. j. o. i. e. Multimodal information presentation: Design guidance and research challenges. **36**, 439–445 (2006).
27. Oviatt, S. & Cohen, P. J. C. o. t. A. Perceptual user interfaces: multimodal interfaces that process what comes naturally. **43**, 45–53 (2000).
28. Newman, E. A. & Hartline, P. H. Integration of visual and infrared information in bimodal neurons in the rattlesnake optic tectum. *Science* **213**, 789–791 (1981).
29. Peck, C. K. Visual-auditory integration in cat superior colliculus: implications for neuronal control of the orienting response. *Prog Brain Res* **112**, 167–177 (1996).
30. Chudler, E. H., Sugiyama, K. & Dong, W. K. Multisensory convergence and integration in the neostriatum and globus pallidus of the rat. *Brain Res* **674**, 33–45 (1995).
31. Olcese, U., Iurilli, G. & Medini, P. Cellular and synaptic architecture of multisensory integration in the mouse neocortex. *Neuron* **79**, 579–593, <https://doi.org/10.1016/j.neuron.2013.06.010> (2013).
32. Wallace, M. T., Meredith, M. A. & Stein, B. E. Multisensory integration in the superior colliculus of the alert cat. *J Neurophysiol* **80**, 1006–1010, <https://doi.org/10.1152/jn.1998.80.2.1006> (1998).
33. Nagy, A., Eördegh, G., Paróczy, Z., Márkus, Z. & Benedek, G. Multisensory integration in the basal ganglia. *Eur J Neurosci* **24**, 917–924, <https://doi.org/10.1111/j.1460-9568.2006.04942.x> (2006).
34. Reig, R. & Silberberg, G. Multisensory integration in the mouse striatum. *Neuron* **83**, 1200–1212, <https://doi.org/10.1016/j.neuron.2014.07.033> (2014).
35. Minciacchi, D., Tassinari, G. & Antonini, A. Visual and somatosensory integration in the anterior ectosylvian cortex of the cat. *Brain Res* **410**, 21–31 (1987).
36. Lee, H., Stirnberg, R., Stocker, T. & Axmacher, N. Audiovisual integration supports face-name associative memory formation. *Cogn Neurosci* **8**, 177–192, <https://doi.org/10.1080/17588928.2017.1327426> (2017).
37. Lanz, F., Moret, V., Rouiller, E. M. & Loquet, G. Multisensory Integration in Non-Human Primates during a Sensory-Motor Task. *Front Hum Neurosci* **7**, 799, <https://doi.org/10.3389/fnhum.2013.00799> (2013).
38. Godfroy-Cooper, M., Sandor, P. M., Miller, J. D. & Welch, R. B. The interaction of vision and audition in two-dimensional space. *Front Neurosci* **9**, 311, <https://doi.org/10.3389/fnins.2015.00311> (2015).
39. Pusztai, A. *et al.* Cortical power-density changes of different frequency bands during a visually guided associative learning test: a human EEG-study. *Frontiers in human neuroscience* **12**, 188 (2018).
40. Quak, M., London, R. E. & Talsma, D. A multisensory perspective of working memory. *Frontiers in human neuroscience* **9**, 197 (2015).
41. Goolkasian, P. & Foos, P. W. Bimodal format effects in working memory. *The American journal of psychology*, 61–78 (2005).
42. Delogu, F., Raffone, A. & Belardinelli, M. O. Semantic encoding in working memory: Is there a (multi) modality effect? *Memory* **17**, 655–663 (2009).
43. Fougny, D. & Marois, R. What limits working memory capacity? Evidence for modality-specific sources to the simultaneous storage of visual and auditory arrays. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **37**, 1329 (2011).
44. Morey, C. C. & Cowan, N. When do visual and verbal memories conflict? The importance of working-memory load and retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **31**, 703 (2005).
45. Calvert, G. A., Campbell, R. & Brammer, M. J. Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Current biology* **10**, 649–657 (2000).
46. Beauchamp, M. S., Lee, K. E., Argall, B. D. & Martin, A. Integration of auditory and visual information about objects in superior temporal sulcus. *Neuron* **41**, 809–823 (2004).
47. Giard, M. H. & Peronnet, F. Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *Journal of cognitive neuroscience* **11**, 473–490 (1999).
48. Myers, C. E. *et al.* Learning and generalization deficits in patients with memory impairments due to anterior communicating artery aneurysm rupture or hypoxic brain injury. *Neuropsychology* **22**, 681 (2008).
49. Hsieh, L. T. & Ranganath, C. Frontal midline theta oscillations during working memory maintenance and episodic encoding and retrieval. *Neuroimage* **85**(Pt 2), 721–729, <https://doi.org/10.1016/j.neuroimage.2013.08.003> (2014).
50. Hsieh, L. T., Ekstrom, A. D. & Ranganath, C. Neural oscillations associated with item and temporal order maintenance in working memory. *J Neurosci* **31**, 10803–10810, <https://doi.org/10.1523/JNEUROSCI.0828-11.2011> (2011).
51. Gevins, A., Smith, M. E., McEvoy, L. & Yu, D. High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. *Cereb Cortex* **7**, 374–385 (1997).
52. Roberts, B. M., Hsieh, L. T. & Ranganath, C. Oscillatory activity during maintenance of spatial and temporal information in working memory. *Neuropsychologia* **51**, 349–357, <https://doi.org/10.1016/j.neuropsychologia.2012.10.009> (2013).
53. Liebe, S., Hoerzer, G. M., Logothetis, N. K. & Rainer, G. Theta coupling between V4 and prefrontal cortex predicts visual short-term memory performance. *Nature neuroscience* **15**, 456 (2012).
54. Lee, H., Simpson, G. V., Logothetis, N. K. & Rainer, G. Phase locking of single neuron activity to theta oscillations during working memory in monkey extrastriate visual cortex. *Neuron* **45**, 147–156 (2005).
55. Worden, M. S., Foxe, J. J., Wang, N. & Simpson, G. V. Anticipatory biasing of visuospatial attention indexed by retinotopically specific-band electroencephalography increases over occipital cortex. *J Neurosci* **20**, 1–6 (2000).

56. Ergenoglu, T. *et al.* Alpha rhythm of the EEG modulates visual detection performance in humans. *Cognitive Brain Research* **20**, 376–383 (2004).
57. Schaefer, R. S., Vlek, R. J. & Desain, P. Music perception and imagery in EEG: Alpha band effects of task and stimulus. *International Journal of Psychophysiology* **82**, 254–259 (2011).
58. Hanslmayr, S. *et al.* Prestimulus oscillations predict visual perception performance between and within subjects. *Neuroimage* **37**, 1465–1473 (2007).
59. Foxe, J. J. & Snyder, A. C. The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention. *Frontiers in psychology* **2**, 154 (2011).
60. Michels, L., Moazami-Goudarzi, M., Jeanmonod, D. & Sarnthein, J. EEG alpha distinguishes between cuneal and precuneal activation in working memory. *Neuroimage* **40**, 1296–1310 (2008).
61. Beck, M. H. *et al.* Short- and long-term dopamine depletion causes enhanced beta oscillations in the cortico-basal ganglia loop of parkinsonian rats. *Experimental neurology* **286**, 124–136, <https://doi.org/10.1016/j.expneurol.2016.10.005> (2016).
62. Stein, E. & Bar-Gad, I. beta oscillations in the cortico-basal ganglia loop during parkinsonism. *Experimental neurology* **245**, 52–59, <https://doi.org/10.1016/j.expneurol.2012.07.023> (2013).
63. Timmermann, L. & Fink, G. R. Pathological network activity in Parkinson's disease: from neural activity and connectivity to causality? *Brain: a journal of neurology* **134**, 332–334, <https://doi.org/10.1093/brain/awq381> (2011).
64. Little, S. *et al.* Controlling Parkinson's disease with adaptive deep brain stimulation. *Journal of visualized experiments: JoVE*, <https://doi.org/10.3791/51403> (2014).
65. Chen, C. C. *et al.* Stimulation of the subthalamic region at 20 Hz slows the development of grip force in Parkinson's disease. *Experimental neurology* **231**, 91–96, <https://doi.org/10.1016/j.expneurol.2011.05.018> (2011).
66. Doya, K. Complementary roles of basal ganglia and cerebellum in learning and motor control. *Current opinion in neurobiology* **10**, 732–739 (2000).
67. Graybiel, A. M. The basal ganglia: learning new tricks and loving it. *Current opinion in neurobiology* **15**, 638–644 (2005).
68. Howard, M. W. *et al.* Gamma oscillations correlate with working memory load in humans. *Cerebral cortex* **13**, 1369–1374 (2003).
69. Ossandón, T. *et al.* Transient Suppression of Broadband Gamma Power in the Default-Mode Network Is Correlated with Task Complexity and Subject Performance. *The Journal of Neuroscience* **31**, 14521–14530, <https://doi.org/10.1523/jneurosci.2483-11.2011> (2011).
70. Chalk, M. *et al.* Attention Reduces Stimulus-Driven Gamma Frequency Oscillations and Spike Field Coherence in V1. *Neuron* **66**, 114–125, <https://doi.org/10.1016/j.neuron.2010.03.013> (2010).
71. Myers, C. E. *et al.* Dissociating hippocampal versus basal ganglia contributions to learning and transfer. *J Cogn Neurosci* **15**, 185–193, <https://doi.org/10.1162/089892903321208123> (2003).
72. Moustafa, A. A., Myers, C. E. & Gluck, M. A. A neurocomputational model of classical conditioning phenomena: a putative role for the hippocampal region in associative learning. *Brain Res* **1276**, 180–195, <https://doi.org/10.1016/j.brainres.2009.04.020> (2009).
73. Smith, Y., Surmeier, D. J., Redgrave, P. & Kimura, M. Thalamic contributions to basal ganglia-related behavioral switching and reinforcement. *Journal of Neuroscience* **31**, 16102–16106 (2011).
74. Koelewijn, T., Bronkhorst, A. & Theeuwes, J. Attention and the multiple stages of multisensory integration: A review of audiovisual studies. *Acta Psychol (Amst)* **134**, 372–384, <https://doi.org/10.1016/j.actpsy.2010.03.010> (2010).
75. Delorme, A. & Makeig, S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods* **134**, 9–21 (2004).
76. Perrin, F., Pernier, J., Bertrand, O. & Echallier, J. Spherical splines for scalp potential and current density mapping. *Electroencephalography and clinical neurophysiology* **72**, 184–187 (1989).
77. Cohen, M. X. *Analyzing neural time series data: theory and practice*. (MIT press, 2014).
78. Parks, N. A., Gannon, M. A., Long, S. M. & Young, M. E. Bootstrap signal-to-noise confidence intervals: an objective method for subject exclusion and quality control in ERP studies. *Frontiers in human neuroscience* **10**, 50 (2016).
79. Ing, A. & Schwarzbauer, C. Cluster size statistic and cluster mass statistic: Two novel methods for identifying changes in functional connectivity between groups or conditions. *PloS one* **9**, e98697 (2014).
80. Puszta, A. *et al.* Cortical Power-Density Changes of Different Frequency Bands in Visually Guided Associative Learning: A Human EEG-Study. *Frontiers in human neuroscience* **12**, 188, <https://doi.org/10.3389/fnhum.2018.00188> (2018).
81. Cohen, M. X. Assessing transient cross-frequency coupling in EEG data. *Journal of neuroscience methods* **168**, 494–499 (2008).

## Acknowledgements

The authors thank to Nóra Cserháti for her help during data collection and all the participants for engaging in the research. This work was supported by Hungarian Brain Research Program Grant KTIA\_13\_NAP-A-I/15 and SZTE ÁOK-KKA Grant No: 2019/270-62-2.

## Author Contributions

A.Pu., B.B., E.G. and A.N. designed the study. A.Pu., D.Ny., Zs.G., X.K., B.B. performed the assessment and documented the findings; A.Pu. analyzed the data; A.Pu., A.Pe. and A.N. organized the study and wrote the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-019-45978-3>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019