Machine learning analysis of cortical activity in visual associative learning tasks with differing stimulus complexity

ÁDÁM KISS[1,4] , KÁLMÁN TÓT[1] , NOÉMI HARCSA-PINTÉR[1],
ZOLTÁN JUHÁSZ[2] , GABRIELLA EÖRDEGH[3], ATTILA NAGY[1] and
ANDRÁS KELEMEN[4*]

[1] Department of Physiology, Faculty of Medicine, University of Szeged, Szeged, Hungary

[2] Department of Electrical Engineering and Information Systems, University of Pannonia, Veszprém, Hungary

[3] Department of Theoretical Health Sciences and Health Management, Faculty of Health Sciences and Social Studies, University of Szeged, Szeged, Hungary

[4] Department of Technical Informatics, University of Szeged, Szeged, Hungary

ABSTRACT

Associative learning tests are cognitive assessments that evaluate the ability of individuals to learn and remember relationships between pairs of stimuli. The Rutgers Acquired Equivalence Test (RAET) is an associative learning test that utilizes images (cartoon faces and colored fish) as stimuli. RAET exists in various versions that differ in the degree of the complexity of the stimuli used in the given version. It has been observed that differences in stimulus complexity can lead to marked differences in test performance, but the related cortical functional differences remain to be elucidated. In the present study, we introduce a Machine Learning- and Independent Component Analysis-based EEG signal processing pipeline, which can detect such differences. RAET and its reduced stimulus complexity variant, Polygon was administered to 32 healthy volunteers and EEG recordings were made with a 64-channel system. The most remarkable differences between RAET and Polygon were detected in the frontal regions, which can be connected to decision making. On the other hand, the parietal regions showed the lowest number of differences between RAET and Polygon. Some task-related activity in the temporo-occipital region was identified, which shows different dynamics depending on visual stimulus complexity.

---

* Corresponding author. H-6720 Szeged, Árpád tér 2. Tel./fax: +36 62 546 396. E-mail: kelemen@inf.u-szeged.hu

INTRODUCTION

Associative learning is a well-known learning mechanism in which different stimuli (antecedents and consequents) are linked to a given stimulus based on a behavioral response. The Rutgers Acquired Equivalence Test (RAET) [1, 2] investigates this learning function with visual stimuli. In the original version of RAET [1], the subjects' task is to pair cartoon faces with colored fish. In a later version called Polygon [3], greyscale circles are paired with polygons. Polygon was created with the aim of studying the effects of reduced stimulus complexity on associative learning, retrieval and generalization.

It was recently demonstrated [3–7] that the complexity of visual stimuli can significantly influence associative learning. This may be because more complex stimuli offer more visual cues, increasing the likelihood of evoking associations, and they may also be easier to verbalize. In the case of the RAET, it must be highlighted that the test uses human faces, which have significant importance in social interactions, they can even be recognized by newborns [8]. Several differences in cognitive performance under RAET-like tests have been observed between different groups recently, but the electrophysiological aspect of these differences remains largely unexplored [9, 10].

In this study, we compared the cortical activity of two equivalence learning tests with machine learning algorithms. Besides the original RAET [1, 2], we also used Polygon [3], which follows the same structure as the original RAET, but utilizes different visual stimuli. While the original RAET uses relatively feature-rich, easily verbalizable stimuli (see Fig. 1), Polygon utilizes two-dimensional geometric shapes that are less complex and more difficult to verbalize.

We aimed to determine how the complexity of visual stimuli affects cortical activity during associative learning. Specifically, we sought to identify the EEG activity of the two learning tests, which are distinguishable. To address this, we applied automated machine learning algorithms to compare the independent components of EEG recordings of patients performing the RAET and the Polygon tests. Machine learning is not new to EEG signal processing; it has been effectively used in single-trial classifications before [11]. This study may advance our understanding of the different psychophysical performances in the RAET and the Polygon tests [3] and could contribute to understand human associative learning in general.

MATERIALS AND METHODS

Participants

The study protocol followed the tenets of the Declaration of Helsinki in all respects and was approved by the Regional Research Ethics Committee for Medical Research at the University of Szeged, Hungary (27/2020-SZTE). Participation was voluntary and without any compensation. Potential volunteers were recruited through the personal networks of the authors. They were

informed about the aims and procedures of the study, and that participation was voluntary and could be terminated at any time without any negative consequences. If they agreed to participate, they provided written informed consent.

Young adults between 18 and 30 years of age were eligible for the study. Exclusion criteria included any neurological or psychiatric condition, or any other condition that could potentially interfere with the volunteer's ability to perform the study tasks, including color blindness.

Altogether thirty-two individuals volunteered, and they were all eligible to participate. However, the data of only 26 of them (14 adult males and 12 adult females) were evaluated due to recording errors. The average age of the volunteers was 23.81 years (with a standard deviation of 5.33 years).

Study procedures

The research was conducted in a quiet, dark room. The learning tests were executed on a personal computer with a cathode-ray tube (CRT) screen (refresh rate: 60 Hz). The participants were sitting at a distance of 57 cm from the computer screen. Regardless of the test (RAET or Polygon), every stimulus fits within a 5 cm by 5 cm square, so when viewed from a given distance, every stimulus appears under the same viewing angle (5° in this case). Each participant was tested individually, no other participants or study personnel were present in the room during the recording. The order of the learning tests was randomized, thus minimizing carry-over effects. To avoid performance anxiety, there was no time limit or forced responses in either of the applied learning tests and there was no time constraint for the individual responses either.

Study protocol

Regardless of the specific test (RAET or Polygon), the test paradigm comprises two main phases: the acquisition phase and the test phase. The tests are divided into trials.

In each trial, the subject's task is to pair an antecedent stimulus with a consequent stimulus. In the acquisition phase, the subject learns the association between the antecedents and the consequents through trial-and-error learning. The computer provides immediate visual feedback on the correctness of the subject's guess (the word "Right" or "Wrong" with a green checkmark or a red X, respectively). The subject must achieve a certain number of consecutive correct answers after the presentation of each new association to be allowed to proceed. This number is 4 when the first association is presented and is increased by 2 upon the presentation of each new association that follows (up to a maximum of 12). Thus, the length of the acquisition phase varies among subjects, depending on how efficiently they learn. There are altogether 4 antecedents and 4 consequents. Out of the 8 possible associations, 6 are taught in the acquisition phase.

In the test phase, no further feedback is given about the correctness of the responses, and the subject must recall the already acquired six associations (retrieval). Two new, hitherto unknown associations are presented as well, where the correct answer can be deduced based on the previously learned associations (generalization or transfer). The retrieval and generalization trials are mixed in the test phase. We excluded the data from the generalization trials in this study because, on the one hand, the number of these trials is low (a total of 12 for one participant), and on the other hand, this process is not closely related to associative learning and its discussion in itself would add up to a separate study. A graphical summary of the possible associations is shown in Fig. 1.

The subjects were instructed to learn the associations between four possible antecedents (A1, A2, B1, B2) and four possible consequents (X1, X2, Y1, Y2) (see table in the Appendix section). The antecedents were cartoon faces in RAET and greyscale circles in Polygon (Fig. 2). The consequents were drawn fish of different colors (red, green, blue, yellow) in RAET and polygons (triangle, square, rhombus and concave deltoid) in Polygon. A more detailed description of the learning tests can be found in [2, 3].

The timing of the learning tests is illustrated in Fig. 3. The moment of the decision for each trial depends entirely on the volunteer and is free from any time pressure. After pressing the button, feedback is displayed for one second. This sequence is repeated during the learning phase. In the test phase, no feedback is given, but the timing remains the same. The selection indicator (green circle) is always shown, regardless of the study phase.

Data acquisition

Simultaneously with the RAET and Polygon tests, the volunteers' EEG waveforms were recorded using a 64-channel Biosemi Active Two device at a sampling rate of 2048 Hz. This system provides a very high isolation mode rejection ratio, hence the expected common-mode interference signal has low power. The Biosemi ActiveTwo EEG device uses active electrodes that amplify the signal within the scalp electrodes to reduce the potential interference from environmental electromagnetic noise sources. This removes the need to use an electrically shielded room for the EEG measurements [12–14]. The headcap and electrode positions remained unchanged between recordings, and the recordings were not interrupted during the tests. The internal filter of the unit had low-pass characteristics with a corner frequency of 400 Hz. The electrode layout used was the standard Biosemi 64-channel configuration (International 10–10 system) [15, 16].

Signal processing and data analysis

The raw recordings from the 64 channels were resampled at 256 Hz after brick-wall low-pass filtering in the Python-MNE 1.6.1 environment [17, 18] using default parameters. Subsequently, the built-in notch filter of MNE was applied at the 50 Hz line frequency and its first harmonic, as well as at the 60 Hz scan frequency of the CRT screen and its first harmonic. A high-pass filter of 1 Hz was used to remove the DC offset and to prepare the data for independent component analysis [19].

After preprocessing, the EEG dataset was segmented into one-second stimulus and response (button press) locked epochs. These epochs were then labeled with the associated event (button press or stimulus appearance), the type of test (RAET or Polygon), and the experiment phase (acquisition or test).

Muscle artifacts were detected in the raw data by their spectral footprint [20]. Epochs containing muscle artifacts were excluded from further analysis.

The remaining epochs were manually inspected for signs of electrode contact errors or artifacts originating from static discharges. These artifacts appear as spikes with ringing on one or more channels. As a result, a few epochs were dropped from each recording. On average, 20 to 60 good epochs remained per recording. Channels with signal fluctuations that were an order of magnitude larger than those of other channels were also excluded during this processing step. Additionally, epochs associated with incorrect answers were excluded by the analysis software. Figure 4 shows a simplified structure of the stored data.

To remove the source mixing effect from the scalp EEG signals, we used the Independent Component Analysis (FastICA [21] algorithm included in the Python-MNE environment). The ICA method can decompose EEG signals into true unmixed sources based on the assumption of mutual independence of non-Gaussian sources. This method is routinely used to detect and remove unwanted ocular, muscle and hear-related artifacts but it can also identify distinct neural sources. Here it is used as a feature reduction technique. Since these components are independent, they can be processed individually and separately, so in every iteration the Machine Learning algorithm will operate on a smaller number of features.

Components representing EEG artifacts (such as blinks) were processed like the others; however, in the end the EEG artifacts were excluded from the results.

Machine learning

To identify differences between the independent components of the two learning tests, we employed machine learning classifier algorithms (Long-Short Term Memory, and Support Vector Classification). The epochs were divided into four groups: learning phase-stimulus appearance, learning phase-button press, test phase-stimulus appearance, and test phase-button press. Each group of epochs was analyzed using four distinct classification algorithms. A component was considered to be different in the Polygon and RAET tests if any of the four classification algorithms detected a difference. Figure 5 visualizes this process.

The following classification algorithms were used. Although we employ two types of machine learning algorithms, four distinct models are generated due to differences in feature selection. Long Short-Term Memory and Support Vector Classification are applied to the raw time-domain signals of the epoched components. Additionally, two more SVC models are trained and evaluated using features derived from the Fourier spectrum of the epochs.

These two Support Vector Classification models differ in their approach to feature extraction: one uses the Fourier-transformed spectrum of the epoch prior to the locking event, and the other uses the spectrum after the locking event. It is important to note that the locking events are positioned in the middle of the epochs. Practically, this means that one model applies the Fourier transformation to the first half of the epoch, while the other applies it to the second half. One of the classifiers used was a 64 layer Long Short-Term Memory (LSTM) neural network [22], configured with a recurrent dropout of 0.2, a dropout of 0.2, and a dense layer with sigmoid activation. The validation split was 20 percent, the batch size was 32, and 70 training epochs were employed during fitting. The model used bias vectors. The activation function was hyperbolic tangent, while the recurrent activation function was sigmoid. The kernel was initialized with the Glorot uniform initializer, while the recurrent initializer was an orthogonal one. The biases were initialized with zeros. The forget gates increased the bias by one at their initialization. None of the kernels, recurrent kernels, output biases vectors were regularized. No kernel or recurrent kernel constraints were implemented. No bias constraints were implemented. The random seed was not initialized. Only the last outputs were considered as the output. The states of the LSTM networks were not stored to help further optimizations. Instead of an unrolled network, a symbolic loop was used as a possible implementation. The remaining hyperparameters were the default values of TensorFlow 2.15.0 [23]. In case of LSTM, the feature variables were the time-series representations of the investigated ICA component for each epoch. The outcome variable was a class label, if the time-series belongs to a RAET or a Polygon test.

Additionally, Support Vector Classification (SVC) [24, 25] with a Radial Basis Function kernel was applied to the time-domain signals as features, with further SVCs used in the frequency domain. In the latter case, the power spectra of the signal are used as features. The spectra of the signal before and after the synchronizing event were analyzed separately. The regularization parameter of the classifier was one. The gamma coefficient for the Radial Basis Function kernel was scaled by the number of features. The classifier used shrinking heuristics. The tolerance stopping threshold was one of one thousand. Having gigabytes of memory on the evaluation computer, a two hundred megabyte cache was provided to the classifier. All classes were equal-weighted. No iteration limit was set. The decision function had a shape like a one-vs-rest decision function. The tie breaking option was not set. No seed was given to the pseudorandom number generator. All the hyperparameters of the SVC algorithm were the default parameters of scikit-learn 1.4.1.post1 [25]. The outcome variable was a class label if the time-series belonged to a RAET or a Polygon.

Each classification was performed on ten different train-test splits, with a test size of 20 percent [26]. Following each fitting, the accuracy distribution was used to estimate the differentiation factor of the component between the RAET and Polygon tests. If the average accuracy of a method exceeded 90 percent, the Independent Component Analysis (ICA) component was plotted.

Human evaluation was performed to determine if the component was due to a channel with a loose contact during recording (indicated by the component involving only one channel). If the component was frontally weighted and the evoked signals contained spikes, it could have been identified as a blink component; however, this scenario was not typical, as the blink components did not distinguish between the learning tests according to the applied learning algorithms. If the component exhibited the same asymmetry as the A/D card of the EEG instrument [20], it was identified as an instrument artifact. As all the other components were considered to be of physiological origin, they were included in the working set of the analysis.

RESULTS

The EEG recordings of 32 participants were collected for this study, with each participant completing both the RAET and Polygon tests. Due to issues with data integrity, high noise levels, or a low number of evaluable epochs, six recordings were excluded from further analysis. Additionally, one recording was dropped because of a high number of artifacts. After these exclusions, EEG recordings from 25 participants were analyzed in detail.

As a result of the described methods, composite plots similar to the one shown in Fig. 6 were generated for every component of each participant if the accuracy of any classifier strategy exceeded 90 percent. This section shows and discusses an example of the generated plots, then summarizes the results of all the plots.

Figure 6 displays an ICA component that shows a difference when using the SVC time-domain classifier. The grouping event was the stimulus appearance in the test phase. The topographic plot indicates that this component is weighted in the frontal region. The evoked potentials of this component exhibit a peak, as shown in the heatmap, in nearly all epochs. However, the timing of this peak differs between the Polygon and RAET tests. The averaged evoked potentials confirm this, as the global maxima occur at different times. The spectrum also
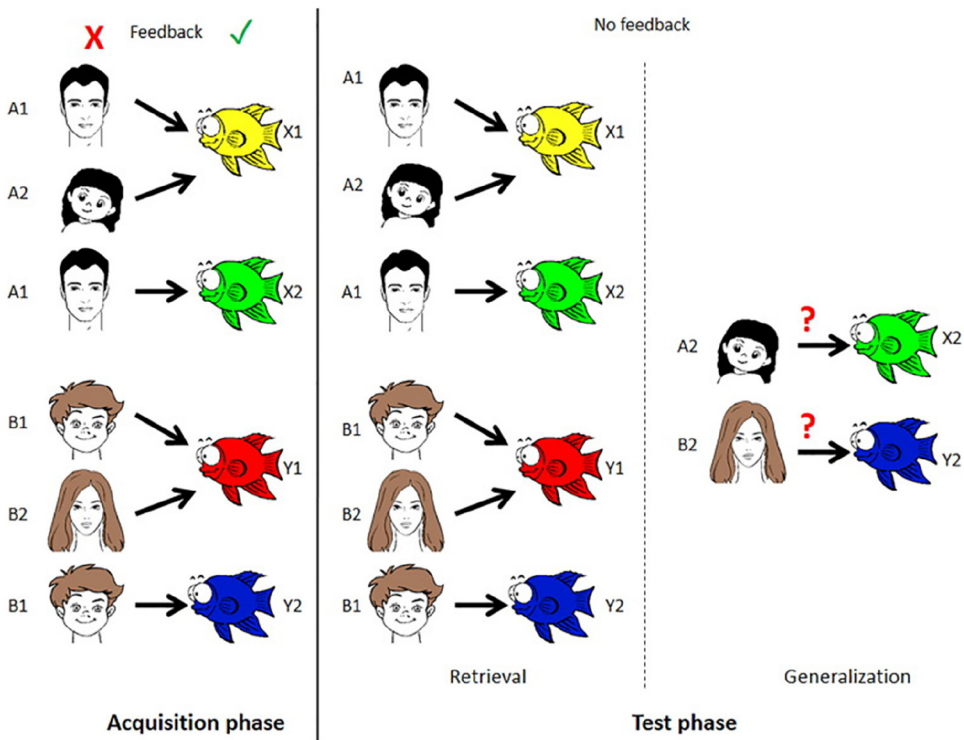
Fig. 1. An overview of the test paradigm with the antecedent–consequent pairs of RAET. The antecedents are cartoon faces of a man (A1) a girl (A2), a boy (B1), and a woman (B2). The consequents are drawings of colored fish: yellow (X1), green (X2), red (Y1), and blue (Y2) [3]

shows a noticeable difference. Examining the variance between epochs does not reveal any outliers. A component like this was categorized as a frontal-weighted component that differentiates the two tests.

Table 1 shows how many recordings/participants resulted in being differentiable by any of the four algorithms. The table clearly shows that the most significant differences between RAET and Polygon were detected in the frontal region. This suggests that the frontal region was the most sensitive to the differences between RAET and Polygon.

On the other hand, the fewest differences were found in the parietal brain area, particularly during the test phase, with only 2 participants showing detectable differences during the "Stimulus appearance" event and 3 during the "Button press" event. This indicates that the parietal region was less sensitive to differences between RAET and Polygon, especially during the test phase.

Overall, the data suggest that the frontal brain area was the most responsive to the differences between RAET and Polygon, while the parietal region showed the least responsiveness during the test phase.

Instead of the accuracy, the area under the Receiver operating characteristic (ROC) curve [11] was considered as a possible metric for the method. In the case of the LSTM, we measured
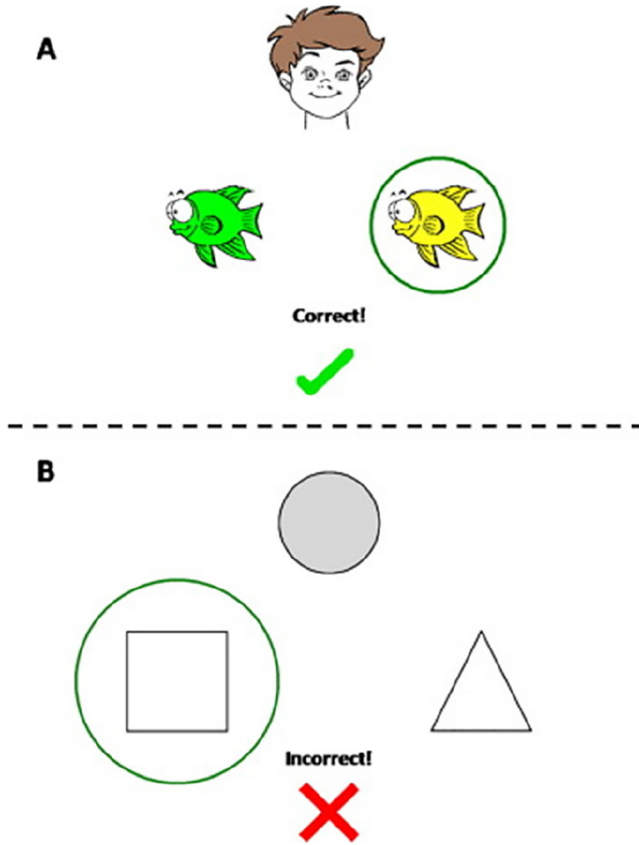
Fig. 2. A trial in the acquisition phase of RAET (A) and Polygon (B). Above is the antecedent and below are the two possible consequents. By selecting the "left" or the "right" fish or polygon, the subject guesses which consequent belongs to the given antecedent. Immediate visual feedback is given. If the guess is right, a green checkmark appears. If the guess is wrong, it is indicated by a red X mark [3]

this area (AUC) [11], and found that it is correlating with the accuracy value. (Correlation coefficient: 0.74.) Most of the outliers are located at very low and very high accuracy or AUC values as shown on Fig. 7. The figure also shows that using the accuracy to determine a threshold is a stronger condition. Furthermore, the AUC value cannot be interpreted the classical way in case of the SVC algorithm. We choose the accuracy as an only metric during the evaluation.

DISCUSSION

To the best of our knowledge, this is the first study that aims to differentiate two RAET-like associative learning tasks with varying stimulus complexity using EEG signals and machine learning. Using the listed results, the comparison of the cortical activity between the original RAET and the feature-reduced Polygon is made possible.
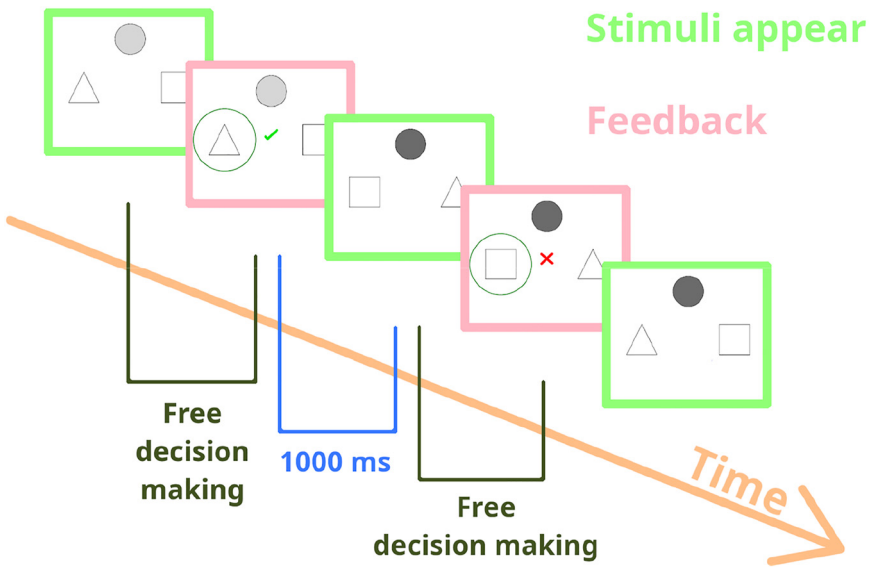
Fig. 3. The timing of the learning tests

The fact that the basal-ganglia take an important role in the memory processes is well-known [27]. However, the basal-ganglia cannot be measured with EEG, only their indirect activity along with the connected brain networks. To investigate the specific role of the inner structures of the brain, functional magnetic resonance imaging (fMRI) can be utilized, with which we can measure the changes in oxygen levels, and estimate the brain activity from it. There are several fMRI studies about cognitive tasks [28]. Using fMRI, it would be possible to deduce if the RAET and the Polygon utilize different brain networks. Comparing such networks is a possible next step in this field.

The EEG signals of 24 out of the 25 evaluated participants (96%) showed differences between the RAET and the Polygon tests. These differences were primarily localized to the frontal region, with the most significant difference observed during the acquisition phase, aligned with the appearance of the stimuli. It is important to note that this timing could overlap with the potential memorization process from the previous trial.

Frontal activity could reflect executive functions [29], attention [30], and working memory load [31]. The only difference between the two tests (RAET vs Polygon) was the complexity (semantic content and verbalizability) of the figures. The differences in cortical activity in the frontal lobe could potentially originate from varying levels of attention, which could affect the decision making. This may result from different learning modes, such as explicit and implicit strategies, including the concealed verbalization of figures. To discuss this appropriately, future research should survey the learning methods used by the volunteers.

The temporal region plays an active role in information processing, language comprehension, and memory processes. Verbalization processes could also affect the results from this region in cases of verbalized learning [32]. To eliminate this possible confounding factor, post-surveying should be conducted in future recordings, as previously mentioned. If we
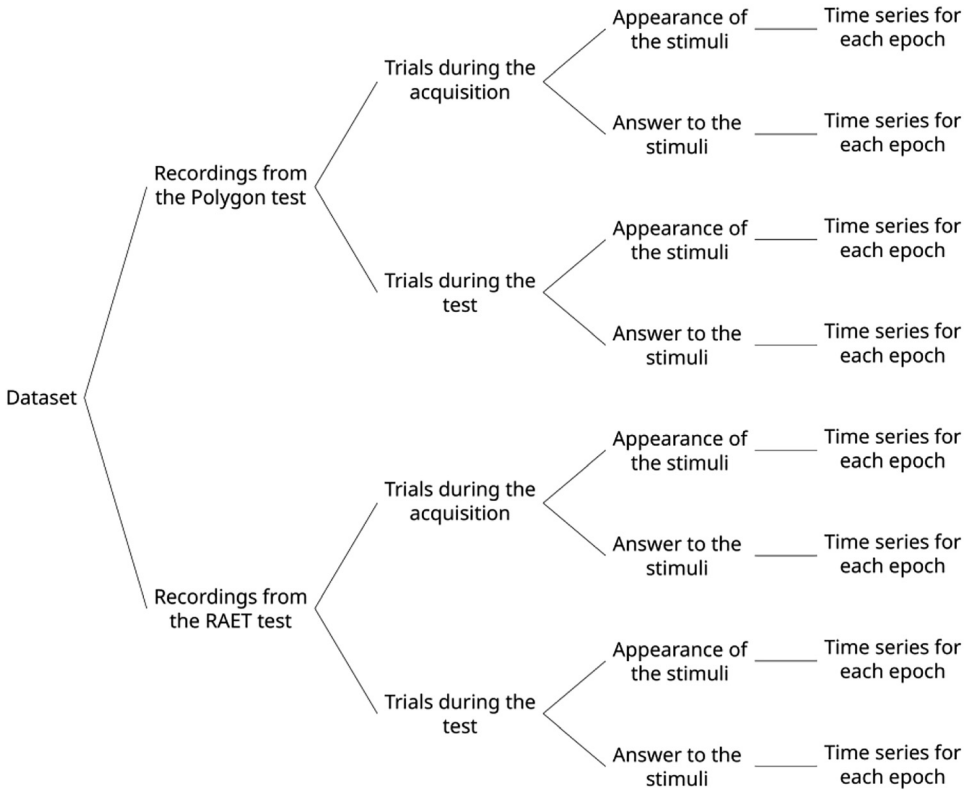
Fig. 4. Organization of the stored data. Every volunteer took both the RAET and the Polygon learning tests. These recordings were separated by the phase of the test, so the acquisition and the test phases are separated. Every trial was further cut into two, aiming to separate the appearance and the answer processes. Some epochs were excluded from the analysis for various reasons discussed in the subsection 'Signal processing and data analysis'
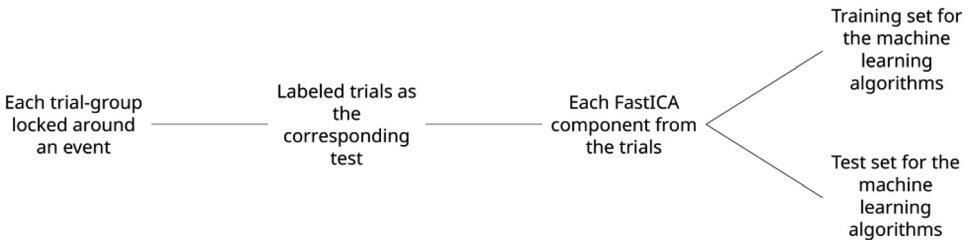
Fig. 5. The data selection pipeline for the machine learning algorithms. The evaluation was done for all the trials groups (learning phase-stimulus appearance, learning phase-button press, test phase-stimulus appearance, and test phase-button press). To utilize the feature reduction property of the FastICA, each component was separately evaluated. For the different trainings, different test-training sets were split
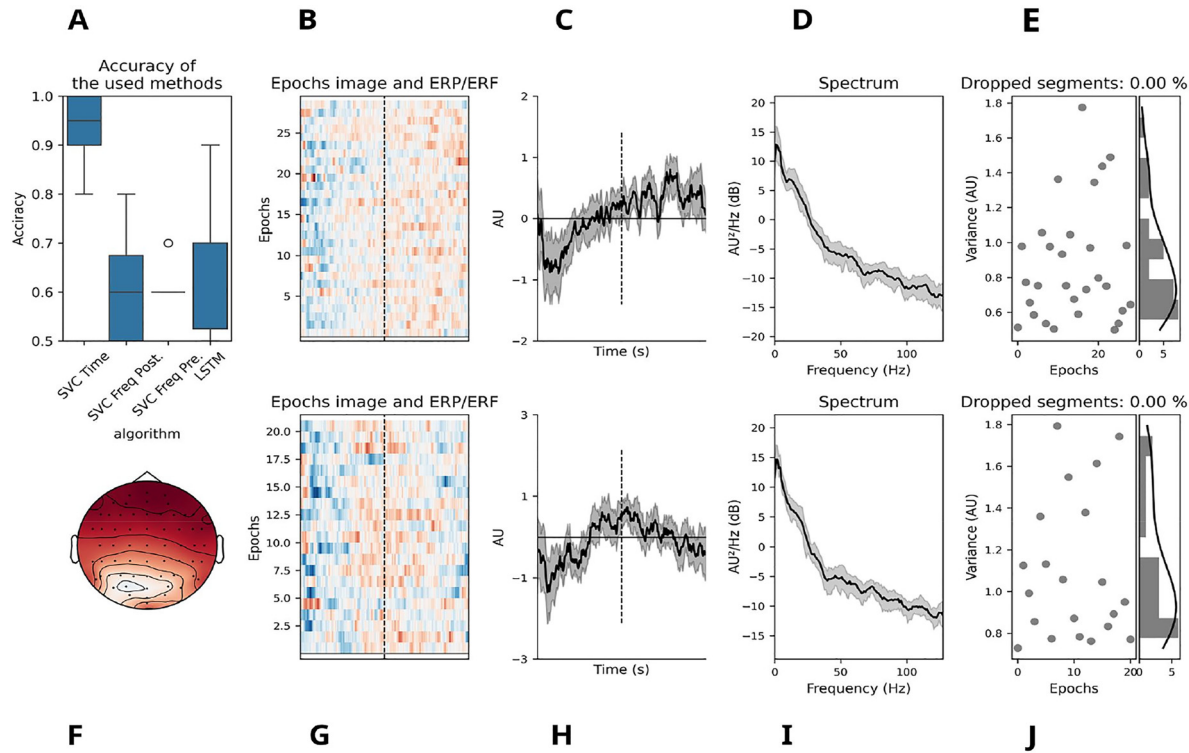
Fig. 6. Example composite plot from the classification algorithms. This ICA component of a patient is frontally weighted and is well distinguishable by the time domain SVC algorithm. The averaged waveform shows a persistent increase in the Polygon test and a decrease in the RAET test. The distribution of this phenomenon is nearly equal, as shown in the B–G subplots. When translating the signal to the frequency domain, a slight elevation differentiates the two learning tests. The processing does not involve outliers, as seen in the variance distribution graphs. The upper plots (B–E) describe the Polygon test, while the lower plots (G–J) describe the RAET test. In the bottom left corner (F), the weights of the ICA component are shown on a topoplot. In the top left figure (A), the accuracies of the different machine learning trainings are plotted. For each learning test, the first image (B, G) shows the epochs on a heatmap. The vertical axes represent epochs and the horizontal axes represent time; Event Related Potential amplitude is represented by color: warmer colors indicate positive, while colder colors indicate negative amplitudes. Plots in the third column (C, H) show the averaged evoked potentials. Plots D and I present the average spectral density of the epochs. The last plots (E, J) show the variance of each epoch. This plot was used as a checkpoint and did not reveal any outliers in the analysis. The dashed line represents the synchronizing event (in this example, the stimulus appearance)

Table 1. Number of identified differences by region and locking event*

	Acquisition – Stimulus appearance	Acquisition – Button press	Test – Stimulus appearance	Test – Button press
Frontal	11 (13)	10 (14)	10 (14)	10 (16)
Occipital	5 (5)	8 (11)	6 (10)	10 (10)
Temporal	7 (9)	7 (10)	5 (13)	5 (7)
Parietal	4 (4)	3 (6)	2 (2)	3 (3)
Overlapping	2 (4)	2 (4)	7 (8)	2 (2)

*The numbers indicate the number of participants for whom differences were detectable over the given brain areas during the corresponding phases of the learning tests at the same events (stimulus appearance or button press). The numbers in parentheses represent the number of ICA components involved. For each person, there are approximately 50 ICA components. Some components were detectable over multiple areas. These are counted in the "Overlapping" row.
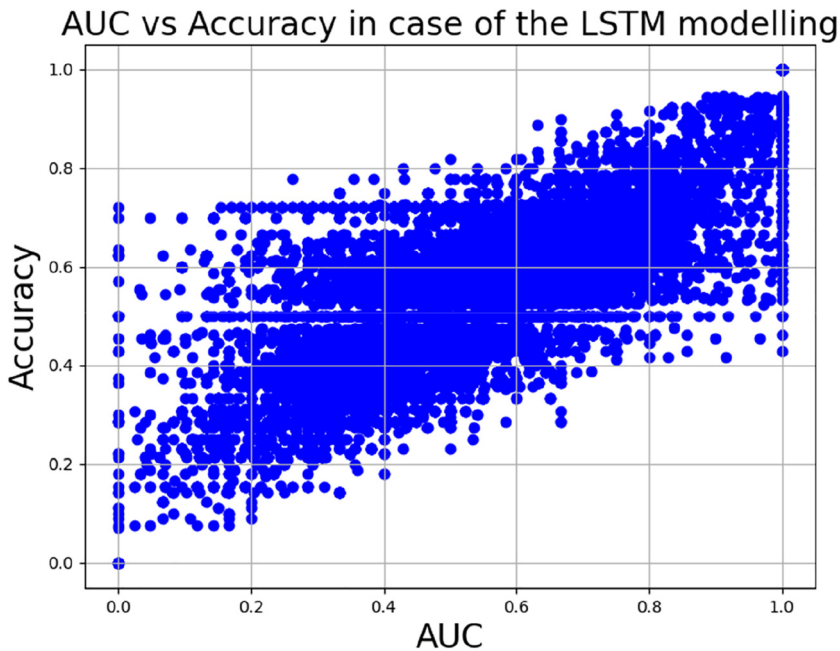

Fig. 7. The relation between AUC and accuracy metrics in the case of the LSTM algorithm

establish that verbalization occurred during the test and influenced the temporal lobe, it would indicate that verbalization differs between the two tests. Since this difference primarily appears in the learning phase, it suggests that the learning process differs in the two scenarios.

The functions of the occipital lobe are connected primarily to visual information processing, which include face recognition [33] and also color determination [34]. It also plays a role in working memory [35] and object recognition [36], and it is connected to the frontal lobe [37].

This area has less differentiability than the frontal lobe, but more deviation between the "Stimulus appearance" and "Button press" events of the study.

Our results indicated that the parietal region was less sensitive to showing activation differences between RAET and Polygon. The parietal regions seem to contribute similarly to the learning and retrieval functions independently from the stimulus complexity and verbalizability.

The main limitation of the study is the small sample size. But the analysis of this limited data set already demonstrated that artificial intelligence could find some stimulus complexity and verbalizability related cortical activity patterns in associative learning of healthy humans. Due to the nature of machine learning, if we need to show the cortical activity differences more precisely, further biomathematical analysis of the EEG data is required.

Statement for conflicts of interest: The authors state that they have no conflict of interest to declare.

Author contribution: Ádám Kiss: Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing – Original Draft, Writing – Review & Editing, Visualization.

Kálmán Tót: Investigation, Data Curation, Writing – Review & Editing.

Noémi Harcsa-Pintér: Writing – Review & Editing.

Zoltán Juhász: Methodology, Validation, Writing – Review & Editing, Visualization.

Gabriella Eördegh: Conceptualization, Investigation, Project administration.

Attila Nagy: Conceptualization, Resources, Writing – Review & Editing, Project administration, Funding acquisition.

András Kelemen: Methodology, Software, Validation, Formal analysis, Writing – Original Draft, Writing – Review & Editing, Visualization, Supervision.

All authors have read and agreed to the published version of the manuscript.

Informed consent statement: Informed consent was obtained from all subjects involved in the study.

Data availability statement: Data from this study can be accessed by contacting the data holder named in the ethical statement, Attila Nagy, by email at nagy.attila.1@med.u-szeged.hu.

ACKNOWLEDGMENTS

REFERENCES

1. Myers CE, Shohamy D, Gluck MA, Grossman S, Kluger A, Ferris S, et al. Dissociating hippocampal versus basal ganglia contributions to learning and transfer. J Cogn Neurosci 2003; 15(2): 185–93. https://doi.org/10.1162/089892903321208123.

2. Öze A, Nagy A, Benedek G, Bodosi B, Kéri S, Pálinkás É, et al. Acquired equivalence and related memory processes in migraine without aura. Cephalalgia 2017; 37(6): 532–40. https://doi.org/10.1177/0333102416651286.

3. Eördegh G, Tót K, Kelemen A, Kiss Á, Bodosi B, Hegedűs A, et al. The influence of stimulus complexity on the effectiveness of visual associative learning. Neuroscience 2022; 487: 26–34. https://doi.org/10.1016/j.neuroscience.2022.01.022.

4. Arslan S, Broc L, Mathy F. Lower verbalizability of visual stimuli modulates differences in estimates of working memory capacity between children with and without developmental language disorders. Autism Dev Lang Impair 2020; 5: 2396941520945519. https://doi.org/10.1177/2396941520945519.

5. Zettersten M, Lupyan G. Finding categories through words: more nameable features improve category learning. Cognition 2020; 196: 104135. https://doi.org/10.1016/j.cognition.2019.104135.

6. Alejandro RJ, Packard PA, Steiger TK, Fuentemilla L, Bunzeck N. Semantic congruence drives long-term memory and similarly affects neural retrieval dynamics in young and older adults. Front Aging Neurosci 2021; 13: 683908. https://doi.org/10.3389/fnagi.2021.683908.

7. Tót K, Eördegh G, Kiss Á, Kelemen A, Braunitzer G, Kéri S, et al. Visual consequent stimulus complexity affects performance in audiovisual associative learning. Sci Rep 2022; 12(1): 17793. https://doi.org/10.1038/s41598-022-22880-z.

8. Simion F, Giorgio ED. Face perception and processing in early infancy: inborn predispositions and developmental changes. Front Psychol 2015; 6: 969. https://doi.org/10.3389/fpsyg.2015.00969.

9. Braunitzer G, Tót K, Eördegh G, Hegedűs A, Kiss Á, Kóbor J, et al. Suboptimal multisensory processing in pediatric migraine without aura: a comparative, cross-sectional study. Sci Rep 2023; 13(1): 19422. https://doi.org/10.1038/s41598-023-46088-x.

10. Rosu A, Tót K, Godó G, Kéri S, Nagy A, Eördegh G. Visually guided equivalence learning in borderline personality disorder. Heliyon 2022; 8(10): e10823. https://doi.org/10.1016/j.heliyon.2022.e10823.

11. Stewart AX, Nuthmann A, Sanguinetti G. Single-trial classification of EEG in a visual object task using ICA and machine learning. J Neurosci Methods 2014; 228: 1–14. https://doi.org/10.1016/j.jneumeth.2014.02.014.

12. MettingVanRijn AC, Kuiper AP, Dankers TE, Grimbergen CA. Low-cost active electrode improves the resolution in biopotential recordings. In: Proceedings of 18th annual international conference of the IEEE engineering in medicine and biology society, 1. Piscataway, NJ: IEEE; 1996. p. 101–2. https://doi.org/10.1109/iembs.1996.656866.

13. Metting van Rijn AC, Peper A, Grimbergen CA. The isolation mode rejection ratio in bioelectric amplifiers. IEEE Trans Biomed Eng 1991; 38(11): 1154–7. https://doi.org/10.1109/10.99079.

14. Metting van Rijn AC, Peper A, Grimbergen CA. High-quality recording of bioelectric events. Part 1. Interference reduction, theory and practice. Med Biol Eng Comput 1990; 28(5): 389–97. https://doi.org/10.1007/BF02441961.

15. BioSemi B.V. Active two specification. [cited 2024 Aug 13]. Available from: https://www.biosemi.com/activetwo_full_specs.htm.

16. BioSemi B.V. Headcaps. [cited 2024 Aug 13]. Available from: https://www.biosemi.com/headcap.htm.

17. Gramfort A, Luessi M, Larson E, Engemann DA, Strohmeier D, Brodbeck C, et al. MEG and EEG data analysis with MNE-Python. Front Neurosci 2013; 7: 267. https://doi.org/10.3389/fnins.2013.00267.

18. Van Rossum G, Drake F. Python/C Api manual-Python 3. Scotts Valley, CA: CreateSpace; 2009.

19. Winkler I, Debener S, Muller K-R, Tangermann M. On the influence of high-pass filtering on ICA-based artifact reduction in EEG-ERP. In: 2015 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC). Piscataway, NJ: IEEE; 2015. p. 4101–5. https://doi.org/10.1109/embc.2015.7319296.

20. Kiss Á, Huszár OM, Bodosi B, Eördegh G, Tót K, Nagy A, et al. Automated preprocessing of 64 channel electroenchephalograms recorded by biosemi instruments. MethodsX 2023; 11: 102378. https://doi.org/10.1016/j.mex.2023.102378.

21. Comon P. Independent component analysis, A new concept? Signal Process. 1994; 36(3): 287–314. https://doi.org/10.1016/0165-1684(94)90029-9.

22. Graves A. Supervised sequence labelling with recurrent neural networks. Berlin, Heidelberg: Springer; 2012. https://doi.org/10.1007/978-3-642-24797-2.

23. TensorFlow Developers. TensorFlow (v2.15.0). Zenodo; 2023. https://doi.org/10.5281/zenodo.10126399.

24. Awad M, Khanna R. Efficient learning machines. Berkeley, CA: Apress; 2015. p. 39–66. Chapter 3, Support Vector Machines for Classification. https://doi.org/10.1007/978-1-4302-5990-9_3.

25. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. J Mach Learn Res 2011; 12: 2825–30.

26. Schaffer C.. Selecting a classification method by cross-validation. Mach Learn 1993; 13(1): 135–43. https://doi.org/10.1007/BF00993106.

27. White NM. Mnemonic functions of the basal ganglia. Curr Opin Neurobiol 1997; 7(2): 164–9. https://doi.org/10.1016/s0959-4388(97)80004-9.

28. Alavash M, Doebler P, Holling H, Thiel CM, Gießing C. Is functional integration of resting state brain networks an unspecific biomarker for working memory performance? Neuroimage 2015; 108: 182–93. https://doi.org/10.1016/j.neuroimage.2014.12.046.

29. Stuss DT. Functions of the frontal lobes: relation to executive functions. J Int Neuropsychol Soc 2011; 17(5): 759–65. https://doi.org/10.1017/S1355617711000695.

30. Chayer C, Freedman M. Frontal lobe functions. Curr Neurol Neurosci Rep 2001; 1(6): 547–52. https://doi.org/10.1007/s11910-001-0060-4.

31. Nissim NR, O'Shea AM, Bryant V, Porges EC, Cohen R, Woods AJ. Frontal structural neural correlates of working memory performance in older adults. Front Aging Neurosci 2017; 8: 328. https://doi.org/10.3389/fnagi.2016.00328.

32. Chengaiyan S, Retnapandian AS, Anandan K. Identification of vowels in consonant-vowel-consonant words from speech imagery based EEG signals. Cogn Neurodyn 2020; 14(1): 1–19. https://doi.org/10.1007/s11571-019-09558-5.

33. Eick CM, Kovács G, Rostalski SM, Röhrig L, Ambrus GG. The occipital face area is causally involved in identity-related visual-semantic associations. Brain Struct Funct 2020; 225(5): 1483–93. https://doi.org/10.1007/s00429-020-02068-9.

34. Critchley M. Acquired anomalies of colour perception of central origin. Brain 1965; 88(4): 711–24. https://doi.org/10.1093/brain/88.4.711.

35. Hallenbeck GE, Sprague TC, Rahmati M, Sreenivasan KK, Curtis CE. Working memory representations in visual cortex mediate distraction effects. Nat Commun 2021; 12(1): 4714. https://doi.org/10.1038/s41467-021-24973-1.

36. Walsh M, Montojo CA, Sheu YS, Marchette SA, Harrison DM, Newsome SD, et al. Object working memory performance depends on microstructure of the frontal-occipital fasciculus. Brain Connect 2011; 1(4): 317–29. https://doi.org/10.1089/brain.2011.0037.

37. Pantazatos SP, Yanagihara TK, Zhang X, Meitzler T, Hirsch J. Frontal-occipital connectivity during visual search. Brain Connect 2012; 2(3): 164–75. https://doi.org/10.1089/brain.2012.0072.

Appendix

Table 2. **A summary of the associative learning tests.** A, B: antecedents (drawn faces in the RAET, greyscaled circles in the Polygon), X, Y: consequents (differently colored fishes in the RAET and simple geometric forms in Polygon tests) [3]

	Acquisition Phase		Test Phase	
Shaping	Equivalence training	New consequents	Retrieval	Generalization
A1 -> X1	A1 -> X1	A1 -> X1	A1 -> X1	
	A2 -> X1	A2 -> X1	A2 -> X1	
		A1 -> X2	A1 -> X2	
				A2 -> X2
B1 -> Y1	B1 -> Y1	B1 -> Y1	B1 -> Y1	
	B2 -> Y1	B2 -> Y1	B2 -> Y1	
		B1 -> Y2	B1 -> Y2	
				B2 -> Y2