Chess Experience and EEG Brain Cortical Organisation: An Analysis Using Entropy, Multivariate Statistics and Loreta Sources

F. T. Rocha\textsuperscript{1}, L. R. Silva Junior\textsuperscript{1}, F. H. G. Cesar\textsuperscript{1}, G. A. Giraldi\textsuperscript{2} and C. E. Thomaz\textsuperscript{1}

\textsuperscript{1}Department of Electrical Engineering, FEI, Sao Paulo, Brazil
\textsuperscript{2}Department of Computer Science, LNCC, Rio de Janeiro, Brazil

Abstract—Chess game has been used as a rich environment to study human cognition and several works in the neuroscientific domain have been done using different brain mapping techniques for this purpose. Here we have processed the electroencephalographical signal to create spatial cognitive brain mappings using entropy, multivariate statistics and Loreta sources. The goal is to disclose the possible differences in the cortical organisation of individuals with different proficiencies during chess problems solving. Volunteers were grouped into two different stages according to their performances, classified as beginners or experienced players. Our experimental results on brain mappings have suggested that both groups recruit visual areas to process the spatial informational of the chess board but beginners may rely more on the linguistic information presented whereas the experienced group seems to count on the executive functions.

I. INTRODUCTION

A chess game involves several cognitive neural computations: object recognition (knowledge of the piece format); spatial movement control according to specific rules (movement permissions of each piece); combination of goal-directed movements (in order to capture a piece); strategy of sequential movement planning (to achieve a check or checkmate); and decision making (to decide for the best move or strategy). Therefore, it’s a rich environment to study neurophysiological dynamics of the neural tissue during complex mental processes. Besides that, chess players achieve different proficiency levels. This allows neuroscientists to investigate not only the neural organisation involved in different cognitive processes but also to understand how the neural network creates news graph topologies as people evolve in the expertise of some mental domain. In neurophysiology, we need a tool that permits volunteers to be comfortably seated in a chess game natural environment to investigate such processes. In this way, electroencephalogram (EEG) may be a good option to acquire brain signals that may be processed and associated to the neural network modulation during the many cognitive processes involved in such game. However, EEG has been neglected to investigate neural spatial organization and functional Magnetic Resonance Imaging (fMRI) has been the preferential tool to explore spatial brain activation.

Now, to override the EEG poor spatial resolution and use it for brain mapping purposes, we propose here to use a well-known technique to disclose the spatial sources of the recorded signal, named Low Resolution Tomography (LORETA) [1]. LORETA uses mean values of EEG or magnetoencephalography (MEG) raw data from several decision makings or events to topologically find the sources of such signals but does not infer any relationship between the located sources. For that, we propose to use a brain mapping technique that involves correlation between electrodes data and entropy calculation to summarize the electroencephalogram signal information recorded by all the electrodes into a single variable for each electrode [2] [3]. With this information - a single entropy value for each electrode - we propose to run factor analysis to disclose the possible covariation between the measured entropy of the electrodes. Finally, by plotting in the same graph both the factor loadings of each extracted factor variables (electrode) and LORETA xyz coordinates, we create the brain mappings showing the possible organisation of the LORETA located sources into neural circuits [2] [3]. The objective here using these techniques is to disclose the possible differences in the cortical organisation of these neural circuits in individuals with different proficiencies during chess problems solving.

The remainder of this paper is described as follows. In the next section, we present a summarized bibliographical review about brain mapping and chess problems solving, using different types of brain signal acquisitions and data processing techniques. In section 3, we present the materials and methods, including the data sample, the protocol, the software created to present the tasks and register the data, and the aforementioned technique to both process the EEG signals and create the brain mappings. Section 4 shows the results of LORETA analysis and the brain mapping technique used here. Section 5 presents the discussion of such results and, finally, in section 6 we give conclusions and future works.

II. PREVIOUS WORK

To the best of our knowledge, there are only three published works using EEG to investigate the neurophysiology of chess game. In 2002, Volke et al. [4] used electroencephalogram and topologically found that the relevant cortical areas of experts during chess problem solving were located rather posterior and more in the right hemisphere if compared with those of the novices. Using event-related potential (ERP), Wright et al. [5] investigated expert’s versus novice’s ERP responses during check situations and piece recognition. It was found that experts showed an enhanced negative ERP component
about 200ms (N2), around FCz and PCz electrodes, with check targets in relation to novices. It infers that larger N2 components reflect matching of current perceptual input to memory, and thus are sensitive to experts’ superior pattern recognition and memory retrieval of chunks. Stepien et al. [6] using nonlinear method, called Higuchi Fractal Dimension, found higher values during the thinking over chess moves than in the players resting state. These works revealed some important facts about the EEG responses to chess reasoning but did not investigate the complexity of cortical organisation in different chess game situations.

On the other hand, using Positron Emission Tomography (PET)-Scan, Nichelli et al. in 1994 [7] were the first ones to describe possible neural circuits involved in different chess problem situations. They calculated the difference in brain activity during 3 different tasks: a spatial localisation task revealed activation on the dorsal occipito-parietal pathway; a piece move identification task revealed activation on lateral, medial and inferior regions of the temporal lobe; a checkmate recognition task revealed activation of the parieto-occipital junction, the left orbito-frontal and right pre-frontal lobe. This work found three equivalences between chess reasoning and already known specific cortical areas functions (i.e. parietal areas and viso-spatial processing; temporal lobe and movement perception; pre-frontal lobe and executive planning).

To investigate the neural differences between groups of individuals that have different chess performances, Amidzic et al. [8], [9] used magnetic imaging (MEG) to compare focal bursts of gamma-band activity. They found that this activity, in amateur chess players, was most evident in the medial temporal lobe, in the region of the perirhinal and entorhinal cortex, hippocampus and related structures. On the other hand, chess grandmasters had more gamma-bursts in the frontal and parietal cortices. This fact led authors to suggest that amateurs may use medial areas to create new memories, whereas experts may use frontal neurons to retrieve information about chunk of pieces stored and analysed by parietal regions. More recently, Bilalic et al. [10] using fMRI showed on experts the enrollment of homologous areas of the right hemisphere during object recognition, besides the common involvement of the left temporal and parietal lateral areas. It inferred that expertise may be the result of a broader network on the brain, involving areas of both hemispheres to process the information about the specific domain of this expertise.

By using fMRI during a resting state, Duan et al. in 2012 [11] found enhanced integration between the caudate nuclei and the default mode network (DMN) in the brain of experts. Next [12], they revealed a broader deactivation of DMN on GM/Ms during chess problem-solving tasks. Together, these findings led the authors to conclude that long-term learning and practice in cognitive expertise may influence large-scale brain networks. DMN deactivation and enhanced functional integration of DMN-caudate circuitry are important neural modulations for a better expert performance. In 2014, Duan et al. [13] examined the overall organisation of brain networks by means of resting-state functional connectivity and graph theoretical analysis. It showed this connectivity was increased in GM/Ms among basal ganglia, thalamus, hippocampus, and several parietal and temporal areas.

From the MEG and fMRI results, we may suppose that experts may use a broader and more specialized neural network to deal with more complex cognitive computations, such as strategy and decision making instead of object recognition. Even when experts have to deal with chess pieces recognition, it seems that they use an occipito-temporal circuit, which has been appointed as a region of expertise in object recognition [14]. It is also a goal of our work to allow such studies of neural organisation by means of EEG signals.

III. MATERIALS AND METHODS

A. Materials

1) The chess software - ChessLab: A software written on C# was created and used to present different chess problem situations on a chessboard, by means of Yes/No questions (Figure 1), which were categorised on 4 types following Volke [4] and 1 type following Nichelli [7] (Table 1). For each category we created and presented to volunteers 10 questions, thus totaling 50 questions. All the questions were randomly grouped and presented to each volunteer one at a time, for as long as they required to read them. They were also informed that after the chessboard was presented they could not read the questions anymore. They pressed the space key after reading each question and next a chessboard having different pieces in different positions was presented. Again, unlimited time was given for volunteers to answer, but now they were advised to be as fast and accurate as they could in order to improve their ranking till the end of the test. They gave their answers by pressing S (for yes) or N (for no) keys and the software presented the next question automatically. It also recorded the response time for each question relating to the beginning of the task.

2) Data sample and Experimental Protocol: We collected data from 28 volunteers but our final data sample was composed of 14 males (only the first and last quartiles) aging between 24 and 56 years old. The EEG signal was registered using 20 electrodes placed according to the 10/20 protocol [15]; impedance below 10 Kohm; band-stop filter 60Hz; sampling frequency of 256 Hz and 16 bits of resolution. Task presentation and EEG recording were synchronized so we could use EEG data from the 2 seconds just before each decision making. We suppose that during these two seconds, volunteers brains were organized to compute the problem and

**TABLE I**

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>Object recognition [4]</td>
</tr>
<tr>
<td>C2</td>
<td>Check [4]</td>
</tr>
<tr>
<td>C3</td>
<td>Checkmate [4]</td>
</tr>
<tr>
<td>C4</td>
<td>Checkmate in one move [4]</td>
</tr>
<tr>
<td>C5</td>
<td>Piece movement [7]</td>
</tr>
</tbody>
</table>
give the answer. All volunteers were students or lecturers from our University.

3) Chess Proficiency: None of our volunteers had the Elo rating [16] (professional ranking system on chess tournaments created by the hungarian Arpad Elo), so we used a performance function defined by Volke [4] to measure each volunteer’s performance. We have chosen this function because it considers not only individual accuracy but also response time of the entire group, balancing this way the final individual rating in relation to the local sample. Additionally, if volunteers answer all the questions with the same choice, attempting to achieve 50% of hits in the least possible time, actually they would score 0. This function is calculated for each volunteer as follows:

\[ H_s = (N_{\text{correct}} - N/2) \times RT_m / RT_s \]  \hspace{1cm} (1)

where \( H_s \) = Honorarium (rating), \( N_{\text{correct}} \) = total amount of correct answers, \( N \) = total amount of questions (50), \( RT_m \) = mean response time of all the volunteers in all questions, \( RT_s \) = mean response time of the volunteer.

B. Methods

Entropy has been used to process and analyse neurophysiological signals from EEG in different domains. One of its uses is to help in diagnosing and characterising neural disorders. It was used to better understand the neural dynamics associated to Alzheimer’s disease [17]. By using it to extract features for a later learning machine classification, it resulted in a good discrimination method for epilepsy [18]. It is also suggested as a useful and discriminative tool to investigate the neuro-dynamic properties of the brain in patients with major depressive disorder during emotional stimulation [19]. Multiscale entropy (MSE) has been used to measure dynamical complexity in physiological systems over a range of temporal scales [20]. Using EEG and MSE, scientists have shown correlation between complexity and creativity among elderly subjects [21].

1) EEG Summarization: Here, we propose the use of entropy, as calculated by Shannon [22], to compute the information provided by each electrode about its connectivity with other ones. Therefore we propose to investigate the covariation among electrode’s entropy value using Factor Analysis. The rationale follows the technique presented by Rocha et al. [23] [24] and used in different cognitive domains such as arithmetic solving and reading process.

EEG records the electrical field potentials generated by the activation of sets of neurons or source signals \( s_l \) located in several distinct cortical areas. The EEG data \( d_i(t) \) recorded at a single electrode \( e_i \) represents a weighted linear sum of underlying source signals over time \( t \), that is:

\[ d_i(t) = \sum_{l=1}^{k} w_l s_l(t). \]  \hspace{1cm} (2)

The weights \( w_l \) are determined by the distance of the cortical source domains \( s_l \) from the electrode pair, the orientation of the cortical patch relative to the electrode pair locations, and the electrical properties of intervening tissues. The number \( k \) and connectivity configuration of active sources are determined by the task being currently processed by the brain.

In order to investigate the possible associations occurring among cortical sources during a specific task, we may use multivariate statistical methods, based on the assumption that these associations may be decoded from the EEG signal recorded during the task solution. But, for that we may summarize the information provided by each electrode \( e_i \), over this time (here it is assumed a time window of two seconds before pressing the Y or N buttons, thus signaling decision-making), about all sources \( s_l \) into a single variable.

Since EEG data are assumed to be a weighted sum of the electrical activity from different sources, correlation analysis of the EEG activity \( d_i(t) \) recorded by the different electrodes \( e_i \) may be used to calculate the entropy information \( h(e_i) \) provided by each electrode \( e_i \) about all \( k \) involved sources \( s_l \) into a single variable [2], [3].

This process can be briefly explained as follows. Given that data \( d_i(t) \) and \( d_j(t) \), furnished by two electrodes \( e_i \) and \( e_j \), provide equivalent information about sources \( s_l \) then the absolute value of correlation coefficient \( c_{i,j} \) calculated for \( d_i(t) \) and \( d_j(t) \) will approach [1], otherwise it will approach 0. The highest uncertainty about the information equivalence provided by \( e_i \) and \( e_j \) occurs when the correlation strength \( c_{i,j} \) approaches [0.5].
Therefore, in the same line of reasoning used by Shannon [22] to define the amount of information provided by a random variable, it is proposed that the informational equivalence \( h(c_{i,j}) \) of \( d_i(t) \) and \( d_j(t) \) furnished by \( e_i \) and \( e_j \) is the expected value \( E(I(c_{i,j})) \) of the information \( I(c_{i,j}) \) provided by \( c_{i,j} \) [2], [3], [25]. Since \( c_{i,j} \) may theoretically assume values between zero and one, we used the \( h(c_{i,j}) \) estimate as calculated by [26]:

\[
h(c_{i,j}) = -[c_{i,j} \log_2(c_{i,j}) + (1 - c_{i,j}) \log_2(1 - c_{i,j})].
\] (3)

Now, given \( q \) electrodes and the average correlation coefficient

\[
\tau_i = \frac{\sum_{j=1}^{q-1} c_{i,j}}{q-1},
\] (4)

the informational equivalence measured by \( \tau_i \) can be written by the following formula

\[
h(\tau_i) = -[\tau_i \log_2(\tau_i) - (1 - \tau_i) \log_2(1 - \tau_i)],
\] (5)

which calculates the information provided by \( d_i(t) \) concerning that provided by all other \( d_j(t) \). Thus,

\[
h(e_i) = \sum_{j=1}^{q-1} \text{abs}(h(\tau_i) - h(c_{i,j}))
\] (6)

computes the information provided by \( d_i(t) \) recorded by \( e_i \) about the sources. In short, in a cognitive task solving, we shall expect:

- if \( c_{i,j} = 1 \) for all \( e_j \) then \( \tau_i = 1 \), \( h(c_{i,j}) = h(\tau_i) \) for all \( e_j \), and consequently \( h(e_i) = 0 \). This indicates that \( d_i(t) \) and the corresponding \( e_i \) do not provide any additional information about the sources \( s_l \);

- if \( c_{i,j} = 0 \) for half of \( e_j \) and \( c_{i,j} = 1 \) for the other half, then \( \tau_i = 0.5 \), \( h(\tau_i) = 1 \), \( h(c_{i,j}) = 0 \) for all \( e_j \), and consequently \( h(e_i) \) is maximum and equal to 1. This indicates that \( d_i(t) \) and the corresponding \( e_i \) discriminate two different groups of electrodes providing information about distinct groups of sources \( s_l \), and

- for all other conditions, i.e. \( 0 < h(e_i) < 1 \), \( h(e_i) \) quantifies the information provided by \( d_i(t) \) about the sources \( s_l \).

Figure 2 shows an illustrative example of an EEG summarisation calculated in the experiments for the C3 electrode.

2) Factor Analysis: We have used Factor Analysis (FA), a well-known multivariate statistical technique, to describe the association among the entropy values of each electrode in a non-supervised and multivariate way [23]. The main idea behind FA is to disclose the correlation relationships among the original variables using a few unobservable random ones, called common factors, to properly represent the data [27].

In particular, let an \( N \times n \) data matrix \( X \) be composed of \( N \) input signals (or trials) with \( n \) variables (or electrodes). This means that each column of matrix \( X \) represents the EEG summarisation of a particular electrode observed all over the

![Figure 2. An illustrative example of an EEG summarisation calculated in the experiments for the C3 electrode. All the calculations have been made using 2 seconds immediately before the decision making.](image)

\( N \) trials. Let this data matrix \( X \) have sample correlation matrix \( R \) with respectively \( P \) and \( \Lambda \) eigenvector and eigenvalue matrices, that is,

\[
P^T R P = \Lambda.
\] (7)

It is a proven result that the set of \( m \) (\( m \leq n \)) eigenvectors of \( R \), which corresponds to the \( m \) largest eigenvalues, minimizes the mean square reconstruction error over all choices of \( m \) orthonormal basis vectors [28]. Such a set of eigenvectors scaled by the square root of the corresponding eigenvalues [27] and calculated as

\[
\hat{\Lambda} = [\sqrt{\lambda_1 p_1}, \sqrt{\lambda_2 p_2}, ..., \sqrt{\lambda_m p_m}]
\] (8)

is known as the factor loadings of the data matrix \( X \) estimated by the principal component method.

The estimated factor loadings \( \hat{L} \) of \( X \) can be rotated in order to improve the understanding of the factors, specially if \( R \) deviates significantly from a diagonal matrix. If \( \hat{L} \) is the \( n \times m \) matrix of estimated factor loadings then

\[
\hat{F} = \hat{L} T
\] (9)

is a \( n \times m \) matrix of rotated estimated factor loadings, where \( T \) is assumed to be an orthonormal \( m \times m \) rotation matrix, that is, \( TT^T = T^T T = I \).

Ideally, we would like to see a pattern of loadings where each subset of electrodes is highly represented by a single factor and has negligible coefficients on the remaining ones, allowing an interpretation of the EEG brain mappings with no overlappings. Thus, our natural choice of the orthonormal matrix \( T \) has been based on the varimax criterion proposed by Kaiser [29].

Therefore, those \( \hat{F} = [\hat{f}_1, \hat{f}_2, ..., \hat{f}_m] \) can then replace the initial \( n \) variables on \( m \) rotated common factor loadings.
These factor loadings would be most expressive in terms of variance information and moreover, the brain mappings would be the most independent ones due to the perpendicular rotation $T$ of the initial factor loadings estimated by the principal components method.

We run factor analysis looking for 3 factors with eigenvalues higher than 1. By plotting the factor loadings (values between 0 and 1) of each variable (EEG channels) using a different color scale for each factor (F1: green to dark blue; F2: yellow to orange; F3: pink to dark red), we create the brain mappings illustrating those channels that may compose possible neural circuits. Then we add the $xyz$ location of the LORETA sources to the same map. This superposition integrates the spatial information from LORETA about the possible sources generating the electroencephalographic signal with the possible association between electrode’s activity as calculated by the entropy/factor analyses technique.

3) LORETA: LORETA (Low Resolution Tomography) uses measurements of scalp electric potential differences (EEG) or extracranial magnetic fields (MEG) to find the 3D distribution of the generating electric neuronal activity with exact zero error localization to point-test sources [1]. LORETA has the capability of identifying 6,430 voxels at 5 mm spatial resolution in cortical gray matter and hippocampus.

IV. RESULTS

A. Behavioural Data

We have calculated our volunteers performance using Volke’s equation [4]. Using the quartile calculation we have grouped them as beginners (7 volunteers, ranging from 24 and 39, mean age 30) with rating $<10$ (first quartile), and experienced players (7 volunteers, ranging from 24 and 56, mean age 31) with rating $>20$ (fourth quartile). The other 18 volunteers were grouped in the second and third quartiles.

B. Brain Mapping - Factor Analysis

The entropy calculation was performed for each volunteer and each decision making. Factor analysis has been carried out by grouping all decision makings depending on volunteer’s proficiency and each one of the five question categories (see section III-A1). Therefore, we generated five brain mappings for each one of the two defined experimental groups (beginner and experienced). Figure 3 illustrates the 20 electrodes positions used in this experiment.

Electrodes P3, Pz and P4 have the highest loadings (dark blue) in Factor F1 computed for both experimental groups. Electrodes C3, Cz and C4 loadings on F1 are greater for the experienced group (E) in comparison to the beginner group (B).

Electrodes F7, F3, Fz and Fp2 have the highest loadings (orange) in Factor F2 computed for group E, whereas loadings of electrodes F4 and F8 are higher in group B if compared to group E. Electrode Fp1 has an important loading for both experimental groups in most of question categories.

The best distinction among experimental groups is provided by Factor F3 (red). It is composed of electrodes F7, T3, T5, O1, O2 and T6 in case of group B. In contrast, electrodes T5, O2 and Oz are the only electrodes loading in Factor F3 for almost all tests in case of group E.

No significant brain mapping differences were found among questions categories for both groups.

C. Brain Mapping - Loreta Sources

Sources located bilaterally at Brodmann Areas (BA) (Figure 6) 7, 19, 39 and 40 are those found nearest to electrodes composing Factor F1 computed for both experimental groups (Figure 7 and Figure 8).
Sources located bilaterally at BA 8, 9, 10 and 11 are found near the electrodes composing Factor F2 for both experimental groups B and E, whereas those located at BA 44, 45 and 46 are associated to Factor F1 more frequently in case of the experienced group E than beginner group B.

Sources located at left BA 22, 37, 38 and 43 are near to Factor F1 electrodes in case of experimental group E, but near to Factor F3 electrodes in case of the experimental group B.

V. DISCUSSION

Factor Analysis is a tool to summarize information about large number of variables associated to a process. The present results clearly show that each multivariate statistical factor Fi - calculated for each experimental group and question category - summarizes information about different sets of sources identified by LORETA analysis. Factor Analysis does not address spatial source location, but it is used here to disclose covariation of source activation that may disclose functional source coupling. In this context Factor Analysis helps to identify those networks recruited to solve a cognitive task.

We have not found any significant brain mapping difference among the questions of different categories as Nichelli et al. (1994) [7] have done. The main reason may be that questions were randomised and mixed and volunteers were not advised about the different kinds of questions. Because of that, their cortical activity may have established a common pattern during the entire task to solve each question. We have not found similarity with the results from Volke et al. (2002) [4] either. But, as Wright et al. (2013) [5] our results allow us to propose that experts superior pattern recognition is related to the matching of current perceptual input to memory retrieval of chunks.

The most discriminant factor, among those 3 with the highest eigenvalues, is Factor F3. In beginners’ Factor 3 we may observe an occipito-temporal association involving areas around T3, T5 and O1 electrodes (Figure 4), summarizing information from sources mostly located at the left temporal lobe. At the same time it is composed of few electrodes in the case of the experienced group.

Our experimental protocol consisted of a two-moment user...
findings indicate that the frontal neurons may be associated with the retrieval, by frontal neurons, about episodic memory [34]. In contrast to the above, sources located at BA 44, 45 and 46 are associated with Factor F2 predominantly in the case of the experienced group. These areas may be considered as part of the executive module of working memory [41] [42], our ability to manage different information at the same time, relating them to each other and seeking for the best possible association for decision making.

VI. Conclusion

The present work has proposed and implemented a computational framework to acquire and process electrophysiological signals associated to chess game, using entropy and statistical methods based on neuroscience and multivariate analysis of data, as well as LORETA sources. We conclude that these results point to an issue already noted in past research on children’s linguistic and mathematic literacy [3] [24]. By using oral or written verbal language to (de)code information about some task, which should be solved mainly by circuits not related to verbal language, such as quantification or calculus, we may hamper the neural circuits establishment that could better solve the task.

Since chess is a cognitive task that involves visuo-spatial processes and time restricted strategic analysis, we may assume that the best neural circuits for its solution should promote a procedural, more implicit, process of information. In fact, the explicit knowledge related to chess, as well as to mathematics, should be presented posterior to visual and motor practice, meanwhile teachers observe student’s evolution in a set of tasks that require as few as possible verbal language for its presentation and answering, emphasizing a learning strategy of pattern recognition.

Historically chess game has been taught empirically, but lacking scientific basis. A better understanding about experienced and non-experienced cortical organisation during chess problems solving may help the development of better teaching methods. We believe that longitudinal brain mapping studies are fundamental to reveal the implicit process by which better players base their reasoning during different moments of their training.

Acknowledgment

The authors would like to thank CAPES and CNPq (444964/2014-2). We would also like to thank Thomas Oliveira and Roberto Goncalves de Magalhaes Junior for their contribution to this research.
running part of the experiment described here. Additionally, we would like to thank the chess teachers Davy D’Israel and Wilson da Silva for the very interesting discussions about chess problems.

REFERENCES


