Libet experiment: Further Analyses & Comparison of Different Modes of Desicion-making

Mojan Izadkhah, Aila Teimoori, Parastoo Azizeddin, and Ali Azizpour

All over history, the human has been eager to find out the nature of free will and has been in doubt about the existence of it. Till the recent decades, there wasn't any scientific method capable of examining this issue. New horizons emerged to study this issue, with the birth of neuroscience and novel methods for tracking brain activity. This paper aims to answer some further questions about this field, using the dataset of "Libet's experiment, its replicability, validity and clinical potential" by M.T. Dominik. This dataset consists of five different experiment modes, containing some special and different modes of the Libet experiment named as W, M, S, P, and Pv series. This article focuses on comparing these various modes of experiments. First, we investigate the distribution of time difference between EMG onset and participant's reported time in W, M, and S tasks, and our results confirm the results in the Libet experiment. We observed that the reported time in the W series was usually before the M series, and in the M series was before the S series. Because of the participants' different accuracy in reporting the exact time, we hypothesize that those participants who reported the time for the W series significantly after their EMG onset must be those whose reported time in the S task was mistaken.(A task is mistaken when the participant reports stimulus time before actual stimulus onset). But our findings did not indicate any correlation between the people with the latent W series and mistaken S series. In another test, RP signals in focused and non-focused modes of decision making (named P and Pv series) were compared, resulting in a significant difference between RP signals' starting time in these two modes, but not in their amplitude. Time and frequency features, which are usually used in classifying EEG signals, of all decision making modes were extracted too. These features were compared, but there wasn't any appropriate feature for separating them.

EMG onset | RP signal | Libet | W | M | S | P | Pv

Introduction

Humans usually find themselves facing different choices during their lives. Some decisions are simple, like the channel to watch on TV, the time to wake up at the weekend, and the food to eat for lunch. On the other hand, some are more complicated, like who to marry and which branch to study.

There exist two main beliefs about the decision-making process. The first one states that every decision is made with free will and full authority, whereas the second one notes that the decisions are due to the unconscious mind, without complete free will. The phrase "free will" has been taken into consideration since thousands of years ago. With the progress of neuroscience during the last centuries, we have been able to investigate the concept of free will, more accurately, relying on brain activity.

One of the pioneering studies on human consciousness was done by Libet [1,2], whose most famous experiment is called the Libet experiment. During this work, subjects were asked to watch a special clock, move their wrists whenever they decided in an arbitrary time, and record the time of each decision [1]. He recognized a particular waveform starting 500ms before each decision, which was discovered by Deecke and Kornhuber previously in 1965 [3]. This waveform is called the readiness potential (RP in short), which is observed in the motor cortex and is followed by muscle movement. This phenomenon made many scientists doubt the existence of free will during decision making.

Some other studies were done by M.T. Dominik [4]. Firstly, they validated the method used by Libet for obtaining the time of the attempt to move [5]. They also provided a set of accurate experiments that were complex forms of Libet experiment with some changes in the design to verify the results of the original experiment. They could verify this in general, but some changes were observed in results' categories [6,7]. Finally, they investigated the possibility of using the Libet experiment in clinical usages. They found some evidence resulting in identifying specific disorders but claimed that the results were Indecisive [4].

We use data in [7] for our research in this paper. This data set consists of five different decision-making modes, each examining the participants' function in a particular situation. Some of these experiments consist of information about arbitrary decision-making; some have information about the participant's accuracy in the proceeder of making a decision, and some limit the participant's choices, so the participant should be focused on making a decision at a particular time. In all of these modes, the participant should report the time that he/she thinks has made the decision. On the other hand, the time of the decision-making procedure is recorded based on the EEG signals too.

Our experiment consists of four main parts. First, similar to the Libet experiment, we will compare the decision-making time participants reported with the recorded time from EEG signals. Second, we will try to find any correlation between the participants' accuracy in one special mode of experiment with the same participants' mistakes during the other decision-making modes. In the next part, modes related to the participants' attention are used to find any relationship or difference between the brain activity in two modes of focused and non-focused decision-making modes [8]. Finally, we will try to extract some well-known time and frequency domain features of EEG signals to separate and classify different decision-making modes to identify which mode of the experiment is a particular EEG signal related to.

This paper has three main parts. In the Methods part, the situations of collecting data during the experiments are explained in detail. Next, in the results part, we will present our hypotheses and explain the result of them, and finally, we will

discuss and conclude our experiments and their results in the Conclusion part.

Materials and Methods

For our research we used a dataset [6,7] consisted of 8 participants(4 males and 4 females all righ-handed) who were asked to perform several tasks. Each task was separated into some distinct trials, each set of 40/41 trails made a series and finally 3 consecutive series made a session. Each session was performed at 1 weak interval. Altogether, they had performed seven sessions with each participant; to be mentioned, the first session was for making participants familiar with the whole process of the tasks.[6,10]

At first, a computer program was loaded on the computer with an introduction page which included personal information about the participant, session ID and the tasks they had to perform in that session. Consecutively an instruction was displayed about their task which differed from one participant to another [4].

They performed different tasks on every patient and in all the experiments, the participants were asked either to make a voluntary movement or to wait for skin stimulus at unknow time to response to it with a click on the left mouse button[4]. They used the following equipment for their experience: an electroencephalograph(EEG) for recording electrical activity of the brain, an electromyograph(EMG) for recording electrical activity of muscles, a skin stimulator and a computer with clock software for collecting introspective data [1,4,9].

In all of their tasks they used Libet's clock which was a cathode-ray oscilloscope (CRO) with a circular regular clock shape display (white dot on black background to maximize the contrast) with 5 inches in diameter. That white dot revolved around the clock and it took 2.56s to completely turn around the clock. During all the experiments the clock was placed at distance of 1.95m from the participants[4].

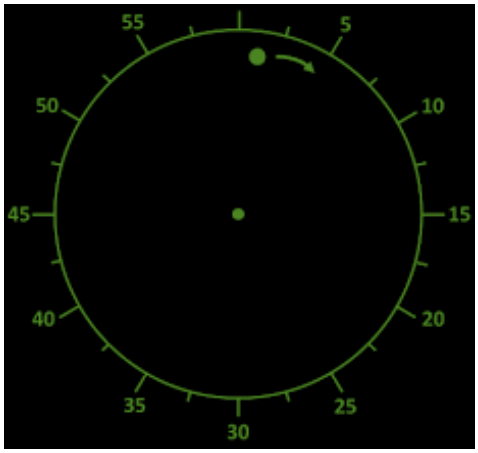

Fig. 1. Libet clock

The purpose of EEG in Libet's experiment is to record readiness potential onsets. For both the EEG and EMG recordings, they used the BIOPAC MP150 unit with EEG100C and EMG100C amplifiers. They recorded the EEG from six standardized 10-20 locations using 0.1 Hz high-pass and 35 Hz low-pass filters. The Fp1, Fp2, Cz, P3, C3 and C4 electrodes were recorded just as what Libet did [9]. The electrodes were embedded in the BIOPAC CAP100C electroencephalography cap filled with an electroconductive gel. For reference electrodes they used both right and left ear lobes. They recorded the EMG from the extensor indicus muscle using two EL503 Ag/AgCl electrodes filled with ELPREP gel. The ground electrode was applied to the upper part of the brachioradialis muscle. They applied 10 Hz high-pass and 500 Hz low-pass filters. Both the EEG and EMG data and the mouse clicks were stored and pre-processed in the AcqKnowledge software.

Another significant instrument in Libet experiment was skin stimulator. They used a non-electrical tactile stimulator TSD190 which was placed on the anterior side of participant's left wrist[4].

Procedure

The subjects sat on a chair and each trial was started whenever the subject claimed that he/she is ready. A tone was played at the beginning of the trial to signal the subject to relax and gaze on the clock in front of him/her. The spot of light on CRO started from 12-clock position and revolving the whole circle for 2.56s. At the beginning of the experiment, subjects were asked to gaze at the center of the CRO and not to follow the spot. Subjects were trained to make self-initiated movements agile[6].

They studied three different kinds of series:

1. Self-initiated voluntary movements (M W Series):

M series consisted of 40 trials in mode A and 41 trials in mode O. As we mentioned before, the participants were sited on a medical armchair, and their right hands were placed on a mouse which was pointing to a fixation point in the center of the Libet's clock. They were asked to click the button whenever they wanted; after clicking the button, they would report the time of beginning the movement(M). After each trial, they could rest their eye if wanted. This experiment was for A and O modes. W series was quite similar to the M series, except that the participants were asked to report the time they were first urged to move [6,4].

2. Skin-stimuli(S Series) :

Like the M series, the S series also consisted of 40 trials in mode A and 41 trials in mode O. This series was mainly made to show how accurate the participants are in time perception. In this event, EMG was not recorded, but instead, a tactile stimulator was placed on the anterior side of the left wrists as they sat on the mentioned chair. The stimuli were produced randomly from the stimulator while the participants were looking at the center of the Libet's clock. After the continuation interval in which they got stimulated, they were asked to recall the time of the stimulation (in this series there were differences between mode A and mode O). Like in the M and W series, the participants were able to rest their eyes between experiments [4].

3. Pre-set motor acts(P Series):

P series, Like the previous series, also consisted of 40 trials. In this series, the sitting arrangement was the same as other

experiments. Participants were again watching the clock's center as their right hand was placed on a mouse (EMG was recorded). In this experiment, in addition to the white circling point, there was also a fixed bright green point. The position of this point was randomly chosen (around the clock) in each trial. The participants were asked to click the mouse button when the white point reached the green point, with the maximum accuracy they can provide. After the mouse click, the green point disappeared while the white point would round for a few more times, then the green point would take a position in a new random place around the clock. In this series, participants did not report the time. If a participant missed recheang of the points, he/she would simply have to wait for the white point to circle again.

The Pv series, in terms of arrangement, setups, and number of trials, were similar to the P series. The only difference was that in this series, the participants were asked to prepare to click the mouse button as the white point reached the green point, but stop their movement before actually clicking the button. If they were successful, they would click the mouse button after that any time they wanted, So the target point got disappeared, and like the previous series, they enter a new trial after the white point circled for a few more rounds [4].

Modes of recall(O & A)

In absolute(A) mode, which was pretty straight forward and simple, the participants were asked to state the clock position when the subjective event happened (in M series, movement or intention to move; in W series, wanting to move and in S series when stimulus delivered). Calculating the mean, they reported the final time in this mode [4,10].

In the order (O) mode, which were more complicated, after the trial finished, the point jumped in an interval, called "stopping range," placed from 400 ms before to 200 ms after the position of the actual stop time. As said, participants were asked to report the time of the relative event, according to their subjective task. The responses could be partitioned into three possibilities [4,10]:

First: reported event occurred before the stop time.

Second: the reported event took place after the stop time

Third: reported event coincides with the exact stop time

According to these types, reported time is calculated with the formula below [4]:

(upper, positive end of "stopping range") – (time interval between "stop times") × (number of points – $\frac{1}{2}$)

Data Analysis

The data was additionally filtered by 0.5Hz high pass filter and 35Hz low pass filter. Three researchers partitioned RPs to three types called I and II and III. we quote their description in the following as these types were detected by eye ball inspection and not by any special program [4,10].

"In type I RP a gradually or steadily rising, ramp-like form begins distinctly prior to -700 ms".

"In type II RPs, the main rise of negativity starts in the range of about -400 to -700 ms, The main portion of this RP is often somewhat dome-shaped rather than ramp-like in form".

"In type III RPs, the main rise of negativity does not appear until about -250 to -200 ms durations of any detectable negativity and especially total areas of RP are also low".

Results

We want to investigate the time of each task in introspective dataset to see the time differences between the tasks. The absolute reports of M(A) and S(A) reports in introspective dataset were gathered from 8 participants, whom performed each task twice for 80 trials and 3 times for W(A) task. As some of the data was missing in some trials, we had to exclude valid ones. The number of excluded valid data in each task is mentioned below.

M(A): 564 from 640 , W(a): 882 from 960 , S(A): 640 from 640

Excluding the response time in respect to EMG onset, we ploted histograms for each task M(M = 31.7004, std = 136.6273), W(M = -92.7846, std = 196.1226) and S(M = 146.0031, std = 150.2124).

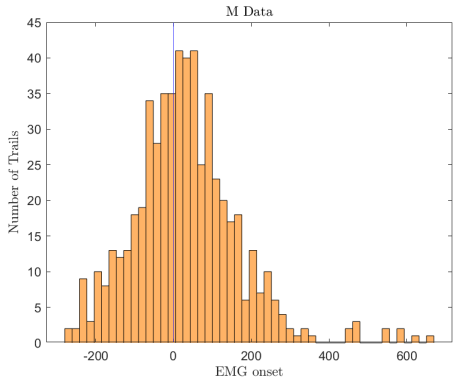

Fig. 2. M introspective data histogram

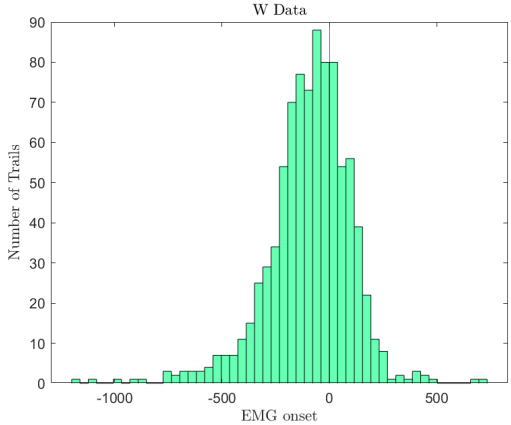

Fig. 3. W introspective data histogram

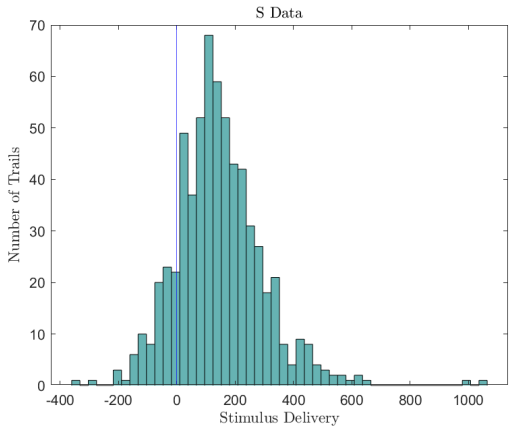

Fig. 4. S introspective data histogram

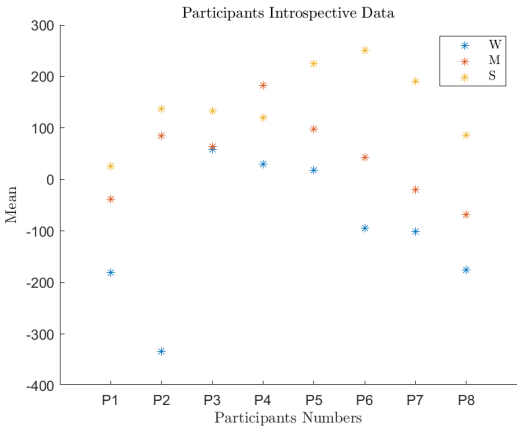

Fig. 6. W & M & S

To demonstrate the reported time of each task, we plotted the 95 % confidence intervals for means of all three tasks in one plot.

Fig. 5. W & M & S

Considering the figures shown above, as we expected there is a considerable difference between times reported as “urge to move” and “the actual movement”. In S series as the reported time is the response time to stimulus, the reported time is after the stimulus in majority of the times.
The figure below illustrates each participant’s W,M and S mean of all times reported in introspective data. As its obvious some of the participants have reported W time after the time that they moved their wrist, means that they state they decided to move their wrist after they actually moved it and that’s wrong. So our alternative hypothesis is that those participants who had reported the time W after they actually moved their wrist, must be those who did not do well in S task (those participants who don’t do well in S task, report times of sensing the stimulus before the actual stimulus come). The reason we had this hypothesis was that as we know some of the participants have data that are not valid and it’s better to be removed from the final analysis in experiment.

For our hypothesis first we calculated mean time of W for each participant and excluded 2 participants who had the largest value of positive mean time because in W series the mean time of all the participant as it’s mentioned is -92.7846s so those participants who have reported the time of decision to move wrist after the actual movement are participants number 3 and 4. According to our hypothesis participants 3 and 4 must have more negative values in their S introspective data ;means that they have reported the time of sensing stimulus before the actual stimulus. But by analyzing the participants number of negative S series , participants number 1 and 8 had the respectively 28 and 13 negative values in S reports that illustrates participants 1 and 8 did not work well in S task which was responsible for checking accuracy of the participants.
The results we get from analyzing these data illustrates the alternate hypothesis was wrong so the null hypothesis which states that tasks are not related and participants can have different performance on different tasks is accepted.
In another test, we decided to examine participants’ attention effect on RP signals. For this purpose, we make a hypothesis as follows: RP signals start earlier when the patient is more focused. So Our null hypothesis is that RP signals’ starting point has no relationship with the participant’s level of focus. To test our null hypothesis, we use the grand average sheet in the EEG.xlsx file. Grand average data is the average of all trials for all participants in six electrodes for five types of experiments. Given that in the P series, the participant is more concentrated on pushing the button at a certain moment, we consider these series of experiments a measure of the focused mode. On the other hand, in the W/M series, the participant is free to push the button whenever he/she wants. So this type of experiment can be a measure of the non-focused mode of the participant.
We plot different electrodes for each experiment to find the RP starting point. Figures 7 and 8 show the EEG average for W/M and P series, Respectively.

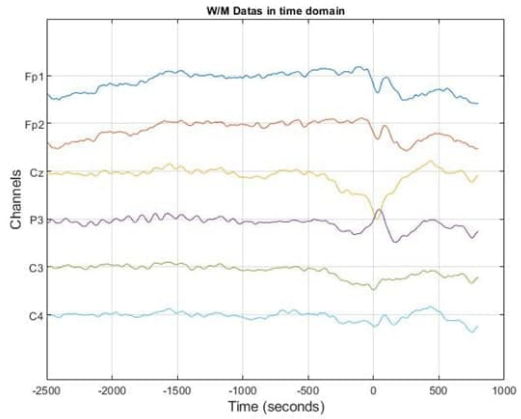

Fig. 7. W/M series Average EEG Signals

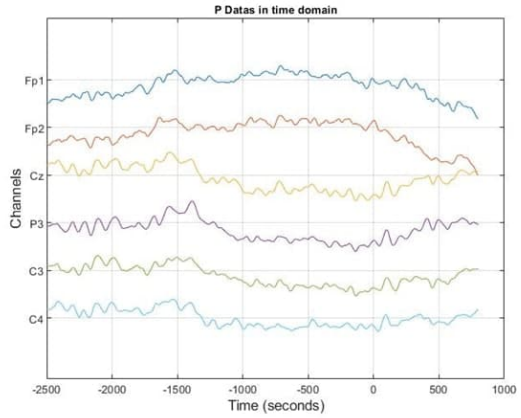

Fig. 8. P series Average EEG Signals

Obviously, Channels Fp1, Fp2, and C4 are not containing considerable RP signals. So we only use channels Cz, P3, and C3 for recording RP signals' starting times, which are obtained for each experiment as follows:

Table 1. RP signals' starting point

Starting point (ms)	Cz	P3	C3
P series	-1396	-1388	-1398
W/M series	-495	-370	-372

Note that the zero time is when the participant makes the decision.

By performing a two-sample t-test on the set of starting points, the obtained p-value is 1.88×10^{-5}. This can reject the null hypothesis, considering that a p-value less than 0.05 states that the null hypothesis can be rejected. So the alternative hypothesis mentioned above is correct, and the RP signals start earlier in focused mode than the arbitrary one. We also compared the amplitude of RP signals in these two series to check whether there are any significant differences between them. So the null hypothesis is that there isn't any difference between the amplitudes of the two series. For this purpose, the minimum amplitudes of the averaged EEG signal during the RP are acquired and shown in table 2.

Table 2. The Rp signals' min & amplitude

Min. amp. (uv)	Cz	P3	C3
P series	-0.01734	-0.02457	-0.031
W/M series	-0.01354	-0.01499	-0.02022

Similarly, a two-sample t-test is performed on the set of amplitudes, resulting in a p-value equal to 0.1437. This states that the null hypothesis can not be rejected so there is no significant difference between the minimum amplitudes of the two modes.

In the next experiment, we tried to compare the two P and Pv modes to determine any noticeable differences between them, helping us recognize whether the participant is willing to push the button or veto.

Similar to the previous experiment, first, the average EEG signals for both series were plotted in the time domain for all six electrodes to observe any difference between the P and Pv series. The time series are shown in Figures 9 and 10.

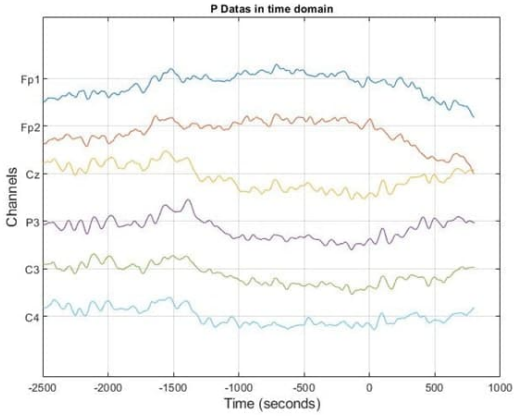

Fig. 9. P series Average EEG Signals

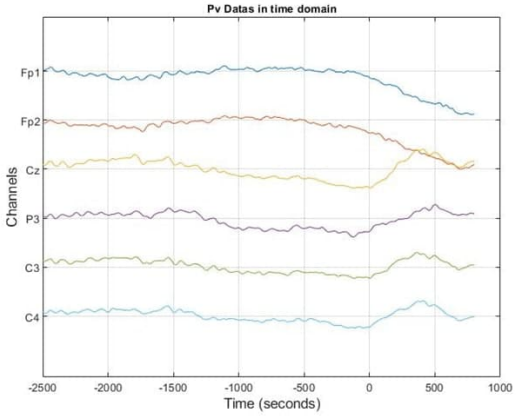

Fig. 10. Pv series Average EEG Signals

As seen in the figures above, although there isn't any considerable difference between the time domain properties such as the starting point and minimum amplitudes of the RP signals in two series, It seems like there may be some difference in the frequency domain. Thus in the next step, the average EEG signals for both series should also be sketched in the frequency domain for all six electrodes, shown in figures 11 and 12.

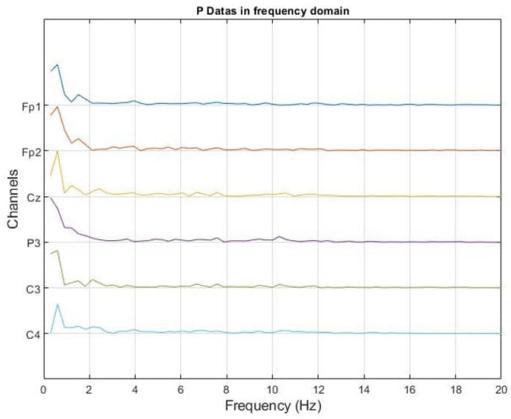

Fig. 11. P series in the frequency domain

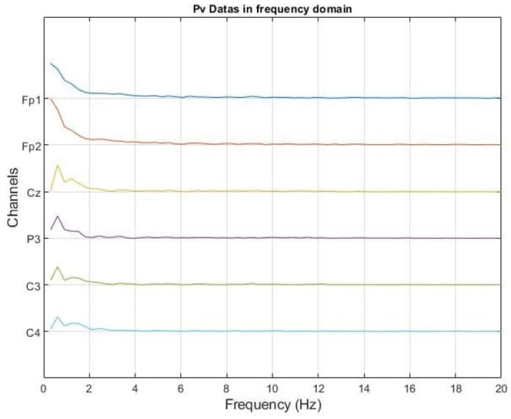

Fig. 12. Pv series in the frequency domain

It's noticeable that despite the similarities in general, there are some observable differences between the P and Pv series in the frequency domain. Hence, some features are extracted from the series to compare them. The features are as follows:
1. The overall signal energy
It is computed by calculating the sum of squared sample values of the signal.

Table 3. The overall signal energy

Channel energy	P series	Pv series
Fp1	18.05	61.06
Fp2	23.75	83.94
Cz	18.85	28.56
P3	19.40	18.85
C3	15.42	13.17
C4	8.32	11.21

Similarly, a two-sample t-test is performed on each mode's values, and the resulting p-value is 0.157, which states that this feature isn't that much different in two series, so It can not separate them efficiently.

From now on, the next features are calculated using the individual-EEG sheet in the EEG.xlsx file. So in the following figures, each point represents each electrode's given feature in each mode. Comparing the two P and Pv sets is done using these features, which were the most common ones when classifying EEG signals.

2. Form Factor
Figure13 shows the distribution of this feature.

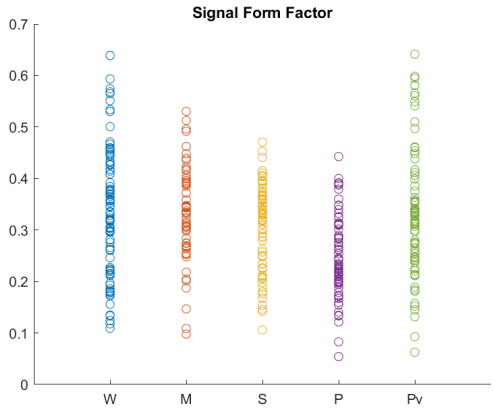

Fig. 13

3. Variance
Figure14 shows the distribution of this feature.

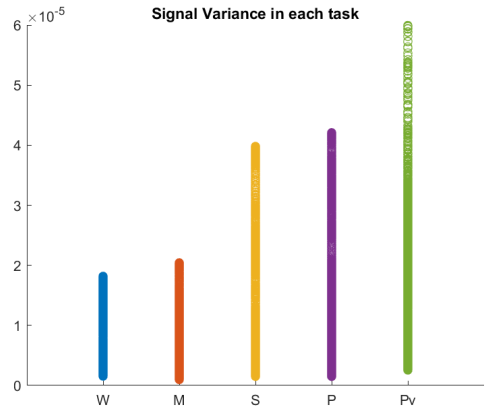

Fig. 14

4. The median and maximum frequency

Figures 15 and 16 show the distributions of these features.

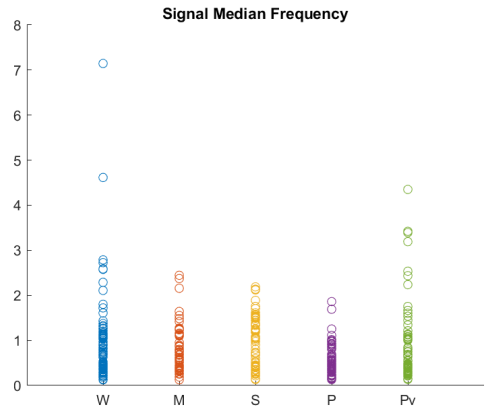

Fig. 15

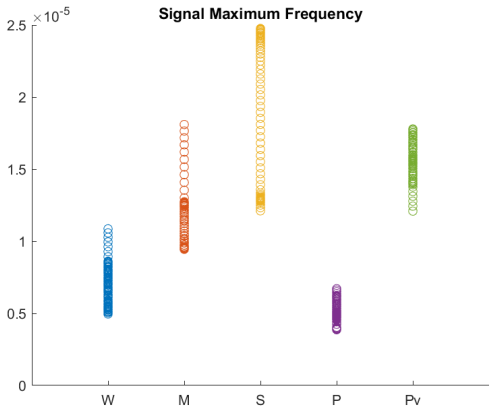

Fig. 16

5. Signal Entropy

Figure 17 shows the distributions of This feature.

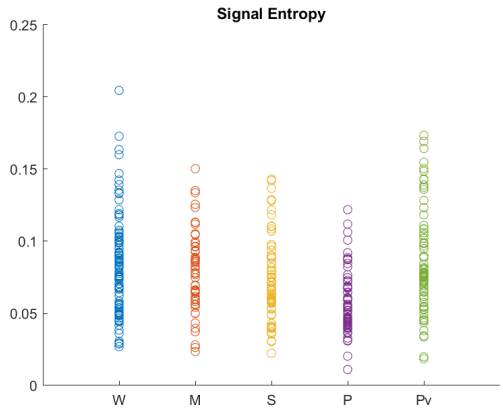

Fig. 17

Finally, there isn't any noticeable difference between P and Pv modes in all features mentioned above, which can separate one from the other. Thus, none of the features 1-5 can perform an efficient method for separating the series from each other to determine whether the participant will push the button or veto.

As seen in figures 13-17, the features 2-5 were calculated for all of the modes (W, M, S, P, and Pv). But similar to the P and Pv modes, other series aren't separable with these features either.

Conclusion

Although our paper aimed to investigate some of Libet's results, it was not our primary goal. Since many research groups have replicated Libet's experiment, we decided to scrutinize Libet's experiment's new horizons. Our first goal was to examine the difference between tasks designed by Libet with new methods. Additionally, we wanted to examine the correlation between each participant's error and mistake in different tasks, but no strong correlation was found.

Furthermore, we found out that the brain activity, more precisely the RP signal, had different features in its time domain in the concentrated and arbitrary decision-making modes, clarifying the fact that our brain acts differently when we are more focused on making decisions. And finally, we aimed to extract some time-frequency features from EEG signals to classify different modes, but there wasn't any feature capable of separating various tasks. Despite the fact that we could not advance previous researches in classifying the tasks, we found it crucial to report them for future studies in this area. More research can be done to complete our findings to classify tasks or find participants' error correlation, but it needs more data collected from different experiments with different methods.

- **Significant Statement**

Decision making is an inseparable part of our lives, and we experience it daily in many situations. Obviously, the brain is the central origin of our decisions and thoughts. So It's clear that it is the organ that should be investigated to discover the mechanism of decision making. In this article, we used brain activities recorded during the decision-making process to discover the nature of this function in the brain. This is important since the appearance of a special waveform in brain activity a few seconds before making decisions can be a sign of an unconscious part interfering with our choices. We also aimed to compare different decision-making modes by comparing the special waveform observed in the brain signals before each decision. Moreover, We hypothesized differences in accuracy between participants and extracted some time-domain and frequency-domain features for classifying different modes in this experiment.

Authors Contribution

Mojan Izadkhah	Parastoo Azizeddin	Aila Teimoori	Ali Azizpour
Designing experiments 1 & 2 & 4	Designing experiments 1 & 2 & 4	Designing experiments 3 & 4	Designing experiments 3 & 4
Materials & Methods	Materials & Methods	Introduction	Introduction
Conclusion & Abstract	Conclusion & Abstract	Conclusion & Abstract	Conclusion & Abstract
Latex	Latex & Significant Statement	Latex Significant Statement	Significant Statement

References

1. Libet B, Gleason CA, Wright EW, Pearl DK. 1983. *Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). The unconscious initiation of a freely voluntary act.* Brain: A Journal of Neurology 106:623.
2. Libet B. 1985. *Unconscious cerebral initiative and the role of conscious will in voluntary action.* Behavioral and Brain Sciences 8:529–539.
3. Kornhuber, H. H., Deecke, L. (1965). *Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale.* Pflügers Archiv, 284, 1–17.
4. Mgr. Tomáš Dominik, 2019. *Libet's experiment, its replicability, validity and clinical potential.*
5. Dominik, T., Dostál, D., Zielina, M., Šmahaj, J., Sedláčková, Z., Procházka, R. (2017). *Libet's Experiment: Questioning the Validity of Measuring the Urge to Move.* Consciousness and Cognition, 49, 255–263.
6. Dominik, T., Dostál, D., Zielina, M., Šmahaj, J., Sedláčková, Z., Procházka, R. (2018a). *Libet's Experiment: A Complex Replication.* Consciousness and Cognition, 65, 1–26.
7. Dominik, T., Dostál, D., Zielina, M., Šmahaj, J., Sedláčková, Z., Procházka, R. (2018b). *EEG Data and Introspective Reports from the Libet's Experiment Replication.* Data in Brief, 20, 2040–2044.
8. Uri Maoz, Gideon Yaffe, Christof Koch, Liad Mudrik,2019.*Neural precursors of decisions that matter—an ERP study of deliberate and arbitrary choice.*
9. Libet, B., Wright, E. W., Gleason, C. A. (1982). *Readiness-potentials preceding unrestricted 'spontaneous' vs. pre-planned voluntary acts.* Electroencephalography and clinical Neurophysiology, 54, 322–335.
10. Libet, Benjamin; Wright Jr., Elwood W.; Feinstein, Bertram; Pearl, Dennis K. (1979). *"Subjective Referral of the Timing for a Conscious Sensory Experience - A Functional Role for the Somatosensory Specific Projection System in Man".* Brain 102: 193–224