

IEEE Copyright Notice

© 2020 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Detection of Careless Mistakes during Programming Learning using a Simple Electroencephalograph

Katsuyuki Umezawa

Dept. of Information Science,
Shonan Institute of Technology
Kanagawa, Japan
omezawa@info.shonan-it.ac.jp

Makoto Nakazawa

Dept. of Industrial Information Science,
Junior College of Aizu
Fukushima, Japan
nakazawa@jc.u-aizu.ac.jp

Manabu Kobayashi

Center for Data Science,
Waseda University
Tokyo, Japan
mkoba@waseda.jp

Yutaka Ishii

Faculty of Education,
Chiba University
Chiba, Japan
yishii@chiba-u.jp

Michiko Nakano

Faculty of Education and Integrated Arts and Sciences,
Waseda University
Tokyo, Japan
nakanom@waseda.jp

Shigeichi Hirasawa

Research Institute for Science and Engineering,
Waseda University
Tokyo, Japan
hira@waseda.jp

Abstract—There are several difficulties encountered by learners during learning such as good or bad learning content, the difficulty level of learning content, and the degree of learning proficiency. It is possible to detect these difficulties by measuring the browsing history, editing history, and biological information such as brain waves or eye-tracking information. In this paper, we measure electroencephalograph (EEG) information during programming learning. We focus on the relationship between task response time and EEG, and try to detect careless mistakes due to the lack of attention. The results show that careless mistakes during programming learning can be detected by experiments.

Index Terms—e-Learning, Self-study System, Simple EEG, Careless Mistake

I. INTRODUCTION

This research aims to develop a self-study system equipped with an artificial teacher who gives advice to students by detecting the learners and to evaluate language learning in a unified framework. “Detecting the learners” means that the system understands the learners’ learning conditions, such as the learner’s degree of understanding, the difference in the learner’s thinking process, the degree of concentration or boredom during learning, and the problem-solving skill of each learner, which can be interpreted from the learning behavior.

Several studies have used electroencephalograph (EEG) information as a way to understand the learners’ learning conditions [1] [2] [3] [4]. It has been found that measuring the ratio of α and β waves is effective for observing the human thinking state [5] [6] [7]. In addition, β/α increases when working on difficult tasks [8] [9] [10].

In this study, we measure EEG information during programming learning. We focus on the relationship between answer time and EEG, and attempt to detect careless mistakes when attention is lost. We believe that intensive training on problems that are likely to cause careless mistakes can lead to improvements in programming skills.

The rest of the paper is organized as follows. In Section 2, we describe previous works related to this study. In Section

3, we explain the details of the experiment. In Section 4, we discuss the proposed method for judging careless mistakes. In Section 5, we analyze the experimental results. Finally, in Section 6, we summarize our study and discuss future work.

II. PREVIOUS WORKS

A. Study on brain waves

It has been empirically found by studies in neuroscience and psychology that brain waves can be used as an index of a person’s state of mind [11]. In previous studies, the α and β waves (obtained from analyzing the power spectra of EEG in frequency space) were used to estimate a person’s state of mind [12].

B. Applicability of brain waves to learning

Several studies have measured cognitive load during learning using EEG. Measuring the changes in the power of brain waves with EEG was also the method of choice in a study that examined cognitive processes in multimedia learning [3]. Antonenko et al. used EEG to investigate the effects of leads on cognitive load and learning in an experiment involving participants reading lead-augmented hypertexts of comparable conceptual difficulty [4].

Other researchers have found that the power spectra of the α and β waves, the ratio of the α and β waves in relation to all brain waves, and the simple ratio of the α and β waves can be used to estimate a person’s state of mind [5] [7]. Moreover, it was also shown that the activity of a person’s brain can be estimated by measuring the α and β waves and estimating β/α [6].

In a previous experiment using a simple typing test with varying degrees of difficulty, we confirmed that the β/α ratio increases with task difficulty [8]. We also found that the low- β -wave/low- α -wave ratio, where “low” means low frequency, increases as a person works on a difficult task [9].

In another experiment in which the change in brain waves was measured as the examinees became used to a new task (assembling a robot using a three-dimensional motion capture), we showed that although the examinees became accustomed to the task in various ways, the low- β /low- α ratio of the examinees who reported that the task was easy fell gradually [10].

III. DETAILS OF THE EXPERIMENT

A. Participants

Three undergraduate students from the Shonan Institute of Technology participated in the experiment.

B. Learning contents

Learning content is the basis of the Java language and consists of the following eight chapters. Note that, Chapter 1 is only for practice; hence, no EEG was measured.

- Chapter 1. Input/output (for practice)
- Chapter 2. Variable, Operation
- Chapter 3. Branch
- Chapter 4. Repetition
- Chapter 5. Array
- Chapter 6. Method
- Chapter 7. Class 1
- Chapter 8. Class 2

Each chapter consists of materials that explain the contents and 10 questions to measure the level of understanding.

Figure 1 shows an example of a question built on Moodle. We conducted a questionnaire on the degree of difficulty that can be answered easily for each question. In addition, since we constructed each question as a single task in Moodle, the answer start time and answer end time can be collected later as logs using the Moodle function.

C. Experiment flow

The flow of the experiment is shown in Fig. 2. Firstly, the participants learn by reading the explanatory material in Chapter 1. Then, they continue to answer questions from the first to the tenth question in Chapter 1. Note again that Chapter 1 is only for practice and the subsequent chapters are the actual experiments. The participants learn by reading the explanatory materials in Chapter 2. The brain wave at that time is measured for 5 minutes, and the measurement is stopped once after 5 minutes. The participants continue to answer questions from the first to the tenth question in Chapter 2. At that time, the brain wave is measured. We then measure continuously without stopping at each question. Since the measurement time can be acquired as a log, it is possible to extract the brain wave corresponding to each question. We repeat the same process for all subsequent chapters.

D. Method of measuring brain waves

The EEG used for measuring the brain waves was a MindWave Mobile headset (NeuroSky, Inc.). As shown in Ref. [15], the MindWave Mobile headset picks up potential differences (voltages) on the skin at the forehead (F_{P1} position

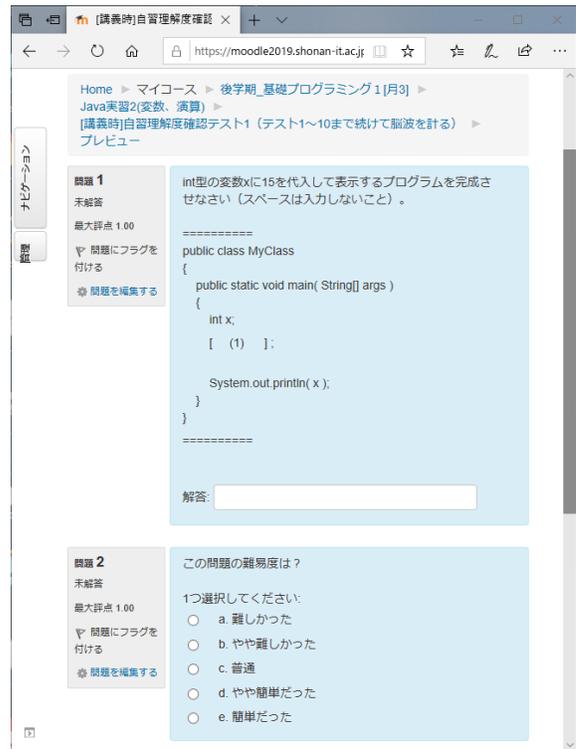


Fig. 1. Question built on Moodle.

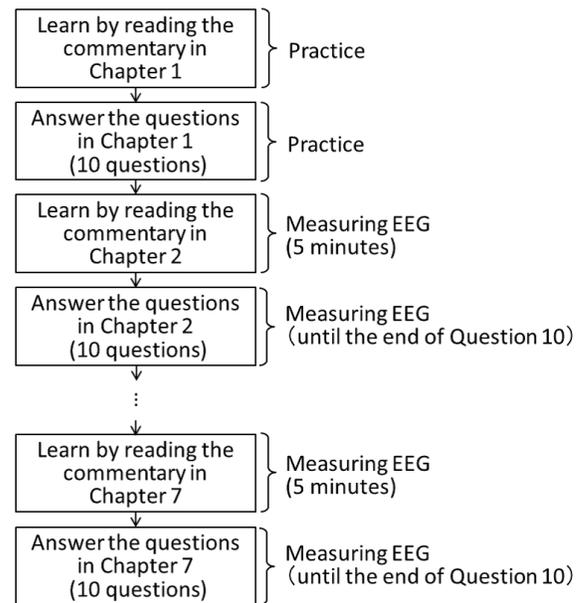


Fig. 2. Experiment flow.

in the international 10-20 EEG system) and the ear (A_1 position)(gray circles in Fig. 3). The signals were passed through analog and digital low- and high-pass filters to retain signals in the 1–50Hz range. After correcting for possible aliasing, the signals were sampled at 128Hz. For each second, the signal was analyzed in the time domain to detect and

correct noise artifacts. A standard fast Fourier transform was performed on the filtered signal, and finally the signal was rechecked for noise and artifacts in the frequency domain.

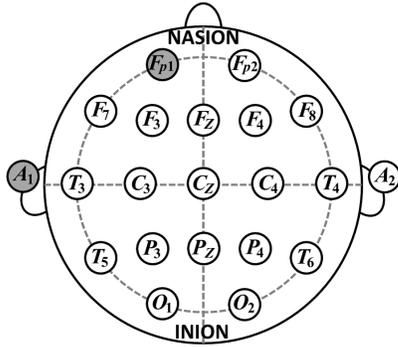


Fig. 3. International 10-20 EEG system.

As shown in Fig. 4, the headset was connected to a ThinkGear Connector by Bluetooth, and the ThinkGear Connector communicated with a log-collection application by TCP/IP. The ThinkGear Connector is a driver provided by NeuroSky Inc., which can transfer brain wave data converted in the headset into the log collection application.

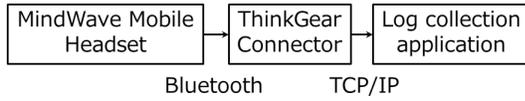


Fig. 4. Outline of the method for measuring brain waves.

Eight types of brain waves were acquired according to Ref. [15]; these are shown in Table I. A four-byte (unit-less) floating-point value was acquired for each type. The MindWave Mobile headset can also collect eSense (i.e., Attention and Meditation) data. The Attention and Meditation values were scaled between 1 and 100, and evaluation of the rating scale was as follows [16]. Measurements between 1 and 20 were considered very low, those between 20 and 40 were considered low, those between 40 and 60 were considered “neutral”, those between 60 and 80 were considered a little high, and those between 80 and 100 were considered high.

TABLE I
ACQUIRED BRAIN WAVES [15].

Type	Frequency (Hz)
δ wave	0.5–2.75
θ wave	3.5–6.75
low α (α_l) wave	7.5–9.25
high α (α_h) wave	10–11.75
low β (β_l) wave	13–16.75
high β (β_h) wave	18–29.75
low γ wave	31–39.75
mid γ wave	41–49.75

IV. PROPOSED CARELESS MISTAKE ESTIMATION METHOD

In this study, we define careless mistakes as “wrong answers in a short time” and “wrong answers without thinking”, that is, “wrong answers because they did not answer properly.”

Our proposed careless mistake estimation method is as follows. First, we calculate the average answer time for each question for each participant in the experiment. Then, we focus on the questions that they answered incorrectly, and then focus on the questions whose answer time is smaller than the average value. In other words, we are focusing on questions answered incorrectly in a short time.

Here, we refer to the average value of the EEG while solving the target question as the target average. We also refer to the average value of the questions from the first question to the question immediately before the target question (not including the target question) as the cumulative average. The next step is to compare the target average with the cumulative average, and extract the value whose target average is smaller than the cumulative average. However, question 1 is excluded because there is no cumulative average value to compare it with. Finally, we performed an F -test (test for homogeneity of variance) and a t -test (test for difference in means) on the extracted average and the corresponding cumulative average, and check whether a significant difference is observed.

V. EXPERIMENTAL RESULT

We first show the average answer time of the participants in Table II.

TABLE II
AVERAGE ANSWER TIME OF THE PARTICIPANTS.

	Average answer time (sec)
Participant 1	39.6
Participant 2	59.6
Participant 3	44.2

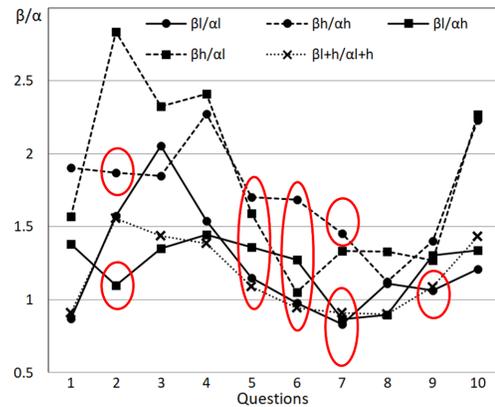


Fig. 5. Average EEG when answering each question in Chapter 3 for participant 1.

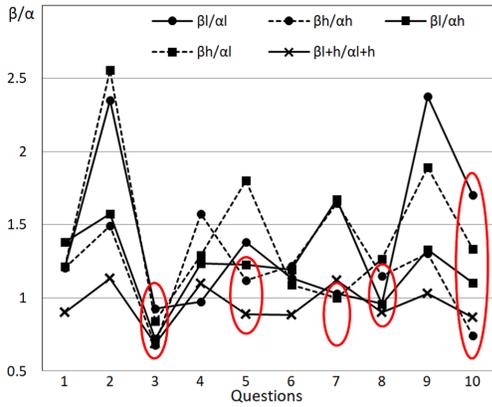


Fig. 6. Average EEG when answering each question in Chapter 7 for participant 2.

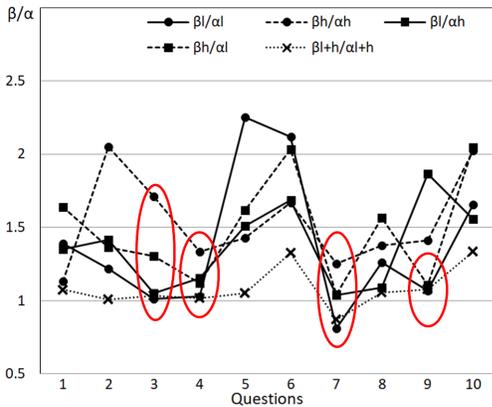


Fig. 7. Average EEG when answering each question in Chapter 8 for Participant 3.

Next, as an example, we show the brain waves of three participants in Figs. 5, 6, and 7. The horizontal axis shows the question number, while the vertical axis shows the value of β/α . Note that we only show the average EEG corresponding to answering each question. We observe that the values of β/α at the locations of the ellipses in Figs. 5, 6, and 7 are decreasing. Hence, if the corresponding questions have incorrect answers, have short answer times, and the values of β/α are statistically significantly lower, then a careless mistake is made.

Table III shows the experimental and estimated results of careless mistakes. It shows the questions that each participant answered incorrectly and the corresponding answer time that was shorter than the average answer time for each participant. We did not include questions that took longer than the average answer time in Table III because they were not considered as careless mistakes.

The tenth to fourteenth columns of Table III show the average values of EEG while answering each question. Note that we considered five types of EEG, i.e., β_l/α_l , β_h/α_h , β_l/α_h , β_h/α_l and β_{l+h}/α_{l+h} , based on a previous study.

The fifth to ninth columns of Table III show cumulative averages. For example, the value of 1.499 in the third row from the top of the fifth column correspond to question 7 in Chapter 3, so it is the average EEG when answering questions 1 to 6 in Chapter 3.

Next, the shaded cells in the tenth to fourteenth columns show values smaller than the cumulative average values in the fifth to ninth columns. We performed an F -test and a t -test on the shaded cells and their corresponding cumulative averages using the original data before averaging (see Appendix). The symbol “*” represents values that are statistically significantly smaller than the average value of EEG.

Cells with this symbol correspond to not only short answer time, but also small β/α from EEG. In other words, it can be determined that they did not consider answering and that this is a careless mistake.

According to Table III, participant 1 had a few incorrect answers. It is thought that he was good at Java programming. Participant 2 had a slightly higher number of incorrect answers, but none were estimated as careless mistakes. It is thought that he made a mistake after careful consideration. Participant 3 made many incorrect answers. Among them were for question 9 in Chapter 2, question 9 in Chapter 4, question 7 in Chapter 5, question 7 in Chapter 6, and questions 3 and 9 in Chapter 8, which can be estimated as careless mistakes. We found that participant 3 was more likely to make careless mistakes at the end of each chapter.

VI. CONCLUSION AND FUTURE WORK

In this study, we measured EEG information during programming learning. We focused on the relationship between question answer time and EEG, and we were able to detect careless mistakes when attention was lost. In the future, we would like to incorporate the method of determining careless mistakes into a self-study system equipped with an artificial teacher who gives advice to students by detecting the learners and evaluate the effectiveness through experiments.

ACKNOWLEDGMENT

Part of this research was conducted as part of the research project “Research on e-learning for next-generation” of the Waseda Research Institute for Science and Engineering, Waseda University. Part of this work was supported by JSPS KAKENHI Grant Numbers JP20K03082, JP19H01721, JP19K04914, and JP17K01101, and Special Account 1010000175806 of the NTT Comprehensive Agreement on Collaborative Research with the Waseda University Research Institute for Science and Engineering. Research leading to this paper was partially supported by the grant as a research working group “ICT and Education” of JASMIN.

REFERENCES

- [1] D. Giannitrapani, “The role of 13-hz activity in mentation,” *The EEG of Mental Activities*, pp. 149–152, (1988).
- [2] P. Antonenko, F. Paas, R. Grabner, and T. van Gog, “Using electroencephalography to measure cognitive load,” *Educational Psychology Review*, Vol. 22, Number 4, pp. 425–438, (2010).

TABLE III
ESTIMATED RESULTS OF CARELESS MISTAKES.

Decision result

Expt. Participants	Chap.	Qs	time (Sec.)	Cumulative averages					Target averages				
				β_l/α_l	β_h/α_h	β_l/α_h	β_h/α_l	β_{l+h}/α_{l+h}	β_l/α_l	β_h/α_h	β_l/α_h	β_h/α_l	β_{l+h}/α_{l+h}
1	3	2	25	0.871	1.904	1.382	1.572	0.908	1.575	1.871	1.097	2.838	1.555
	7	3	29	0.709	0.982	0.778	0.737	0.570	0.734	0.800	1.083	0.561	0.560
2	3	7	39	1.499	1.351	1.385	1.302	0.976	1.146	1.518	1.959	1.093	0.993
	4	8	17	1.038	1.257	1.274	1.002	0.856	1.128	0.946	1.303	0.829	0.880
	5	10	27	1.372	1.240	1.303	1.292	0.950	1.273	1.216	1.380	1.133	0.964
	7	5	38	1.167	1.218	1.216	1.254	0.928	1.380	1.122	1.226	1.800	0.887
	7	7	30	1.180	1.208	1.211	1.264	0.911	1.028	1.649	1.672	1.000	1.120
	7	8	35	1.168	1.243	1.248	1.243	0.928	0.962	1.149	0.956	1.267	0.899
	8	2	28	1.316	1.369	1.259	1.176	1.068	0.860	1.197	1.176	0.841	0.871
3	2	6	40	1.636	1.811	1.375	1.967	1.239	1.199	2.130	1.495	1.834	1.142
	2	7	24	1.583	1.849	1.389	1.951	1.228	1.178	2.198	1.166	1.822	1.202
	2	9	39	1.469	1.761	1.386	1.781	1.172	1.020	1.290*	1.365	0.992*	1.017
	3	9	11	1.455	1.673	1.338	1.673	1.151	1.412	1.387	1.096	2.160	1.262
	3	10	34	1.453	1.662	1.329	1.691	1.155	1.172	1.965	1.450	1.638	1.156
	4	3	26	1.283	1.330	1.201	1.638	1.013	1.376	1.488	1.111	1.901	1.158
	4	5	14	1.266	1.477	1.154	1.754	1.112	2.581	2.232	1.620	2.445	1.488
	4	9	43	1.370	1.461	1.121	1.759	1.101	0.899*	1.275	1.057	1.165*	0.926
	5	7	40	1.349	1.410	1.337	1.458	1.071	0.904*	1.536	1.110	1.229	1.025
	5	10	26	1.346	1.454	1.356	1.507	1.098	2.268	1.448	1.291	2.218	1.342
	6	3	43	1.283	1.701	1.392	1.468	1.033	1.010	1.712	1.053	1.305	1.034
	6	5	24	1.144	1.601	1.241	1.330	1.029	2.252	1.430	1.511	1.616	1.052
	6	7	38	1.488	1.602	1.376	1.532	1.105	0.812*	1.251	1.040	1.044*	0.871*
	6	9	17	1.377	1.525	1.289	1.488	1.072	1.067	1.413	1.864	1.105	1.076
	7	3	30	1.197	1.449	1.191	1.541	1.093	1.279	1.658	1.216	1.540	1.118
	7	4	33	1.222	1.514	1.199	1.540	1.101	1.731	1.551	1.381	2.377	1.223
	7	5	36	1.356	1.524	1.247	1.760	1.133	1.244	1.323	1.167	1.286	0.994
	7	9	40	1.289	1.643	1.317	1.747	1.169	2.926	1.729	1.605	2.426	1.569
8	3	18	1.648	1.810	1.363	2.202	1.216	2.918	1.018*	1.116	1.715	0.996	
8	9	30	1.968	1.727	1.408	2.603	1.266	1.106*	1.714	1.134	1.560*	1.226	

- [3] I. Gerlic, and N. Jausovec, "Multimedia: Differences in cognitive processes observed with EEG," Educational Technology Research and Development, 47(3), pp. 5–14, (1999).
- [4] P. Antonenko, and D. Niederhauser, "The effects of leads on cognitive load and learning in a hypertext environment," Computers in Human Behavior, 26, pp. 140–150, (2010).
- [5] H. Uwano, K. Ishida, Y. Matsuda, S. Fukushima, N. Nakamichi, M. Ohira, K. Matsumoto, and Y. Okada, "Evaluation of software usability using electroencephalogram - comparison of frequency component between different software versions," Journal of Human Interface Society, 10(2), pp. 233–242, (2008). (in Japanese)
- [6] K. Yoshida, Y. Sakamoto, I. Miyaji, and K. Yamada, "Analysis comparison of brain waves at the learning status by simple electroencephalography," The Institute of Electronics, Information and Communication Engineers (IEICE) Technical report, ET, 112(224), pp. 37–42, (2012). (in Japanese)
- [7] K. Yoshida, Y. Sakamoto, I. Miyaji, and K. Yamada, "Analysis comparison of brain waves at the learning status by simple electroencephalography," KES'2012, Proceedings, Knowledge-Based Intelligent Information and Engineering Systems, pp. 1817–1826, (2012).
- [8] K. Umezawa, T. Ishida, T. Saito, M. Nakazawa, and S. Hirasawa, "Collection and analysis of the browsing history, editing history, and biological information for high school students," National Conference of JASMIN 2016 Autumn, Japan Society for Management Information, D2-1. pp. 1–6, (2016). (in Japanese)
- [9] K. Umezawa, T. Ishida, T. Saito, M. Nakazawa, and S. Hirasawa, "A judgment method of difficulty of task for a learner using simple electroencephalograph," Information Processing Society of Japan (IPJSJ) SIG Technical Report, (2016). (in Japanese)
- [10] K. Umezawa, T. Ishida, T. Saito, M. Nakazawa, and S. Hirasawa, "A study on transition of the brain waves when executing a task using simple electroencephalograph," Information Processing Society of Japan (IPJSJ) SIG Technical Report, (2017). (in Japanese)
- [11] N. Yoshimine, "Mystery of the brain: Why the brain-waves oscillate rhythmically?," The Bulletin of Tama University, pp. 93–100, (2017).
- [12] H. Berger "On the electroencephalogram in man," Archiv fur Psychiatrie and Nervenkrankheiten, 87, pp. 527–570, (1929).
- [13] F. Hirai, K. Yoshida, and I. Miyaji, "Comparison analysis of the thought and the memory at the learning time by the simple electroencephalograph," Multimedia, Distributed, Cooperative, and Mobile Symposium (DICOMO 2013), pp. 1441–1446, (2013)
- [14] F. Hirai, K. Yoshida, and I. Miyaji, "Trial of the EEG state feed-back learning system at the time of the memory work by the simple electro-

TABLE IV
RESULTS OF THE F -TEST AND t -TEST.

Expt. Participants	Chap.	Qs	Time (Sec.)	p -value of F -test					p -value of t -test				
				β_l/α_l	β_h/α_h	β_l/α_h	β_h/α_l	β_{l+h}/α_{l+h}	β_l/α_l	β_h/α_h	β_l/α_h	β_h/α_l	β_{l+h}/α_{l+h}
1	3	2	25	–	0.185	0.001	–	–	–	0.948	0.456	–	–
	7	3	29	–	0.002	–	0.009	0.204	–	0.436	–	0.282	0.927
2	3	7	39	0.000	–	–	0.000	–	0.108	–	–	0.228	–
	4	8	17	–	0.001	–	0.001	–	–	0.131	–	0.309	–
	5	10	27	0.000	0.005	–	0.000	–	0.746	0.882	–	0.561	–
	7	5	38	–	0.412	–	–	0.001	–	0.973	–	–	0.638
	7	7	30	0.000	–	–	0.000	–	0.430	–	–	0.172	–
	7	8	35	0.000	0.046	0.000	–	0.137	0.198	0.666	0.104	–	0.802
3	8	2	28	0.000	0.011	0.180	0.004	0.008	0.075	0.593	0.716	0.116	0.246
	2	6	40	0.000	–	–	0.003	0.000	0.217	–	–	0.785	0.462
	2	7	24	0.000	–	0.034	0.006	0.377	0.280	–	0.402	0.810	0.922
	2	9	39	0.000	0.000	0.273	0.000	0.023	0.052	0.006	0.941	0.000	0.294
	3	9	11	0.001	0.013	0.187	–	–	0.881	0.466	0.498	–	–
	3	10	34	0.000	–	–	0.100	–	0.282	–	–	0.903	–
	4	3	26	–	–	0.247	–	–	–	–	0.753	–	–
	4	5	14	–	–	–	–	–	–	–	–	–	–
	4	9	43	0.000	0.000	0.099	0.000	0.000	0.009	0.310	0.715	0.004	0.095
	5	7	40	0.000	–	0.000	0.000	0.000	0.004	–	0.311	0.243	0.670
	5	10	26	–	0.310	0.034	–	–	–	0.986	0.825	–	–
	6	3	43	0.001	–	0.000	0.001	–	0.213	–	0.125	0.594	–
	6	5	24	–	0.039	–	–	–	–	0.757	–	–	–
	6	7	38	0.000	0.000	0.000	0.000	0.000	0.001	0.164	0.087	0.017	0.025
6	9	17	0.001	0.000	–	0.000	–	0.228	0.705	–	0.118	–	
7	3	30	–	–	–	0.127	–	–	–	–	1.000	–	
7	4	33	–	–	–	–	–	–	–	–	–	–	
7	5	36	0.146	0.411	0.011	0.008	0.023	0.743	0.509	0.698	0.231	0.491	
7	9	40	–	–	–	–	–	–	–	–	–	–	
8	3	18	–	0.030	0.105	0.016	0.110	–	0.043	0.439	0.390	0.246	
8	9	30	0.000	0.062	0.000	0.000	0.141	0.023	0.966	0.159	0.015	0.827	

encephalograph,” Multimedia, Distributed, Cooperative, and Mobile Symposium (DICOMO 2014), pp. 633–638, (2014).

[15] ThinkGear Serial Stream Guide, http://developer.neurosky.com/docs/doku.php?id=thinkgear_communications_protocol (2020/1/16 accessed)

[16] MindWave Mobile: User Guide August 5, 2015.

[17] Y. Nagata, “Introduction Statistical Analysis,” JUSE Press. Ltd. (1992).

According to the results of the F -test, the results of the t -test assuming equal variances and assuming not equal variances are shown from the tenth to fourteenth columns of Table IV. The dark shaded cells denote that the p -value is less than 0.05, that is, there is a difference in the average value.

APPENDIX

Table IV shows the results of the F -test and t -test. The results of the F -test are shown in the fifth to ninth columns of Table IV. The unshaded cells denote that the p -value is less than 0.2, that is, the null hypothesis of the F -test (test if two populations have equal variance) is rejected and the two populations are determined to not have equal variance. Conversely, cells that are lightly shaded denote that the two populations have equal variance.