Differential analysis of heart rate, facial expressions and brain wave during learning of visual- and text-based languages

Katsuyuki Umezawa
Department of Informatics,
Shonan Institute of Technology
Kanagawa, Japan
umezawa@info.shonan-it.ac.jp

Takumi Koshikawa
Department of Information Science,
Shonan Institute of Technology
Kanagawa, Japan
19A3096@sit.shonan-it.ac.jp

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

Shigeichi Hirasawa
Data Science Center,
Waseda University
Tokyo, Japan
hira@waseda.jp

***Abstract*—From 2020, programming education has become compulsory in elementary schools. Visual-based programming languages are becoming popular as an introduction to programming. At universities, students learn text-based programming languages such as C and Java. Much research has been conducted on the transition from visual- to text-based programming languages. However, most related studies are limited to questionnaire evaluation after learning. In this study, we focus on evaluation during learning. Specifically, 18 types of biometric information (heart rate, 12 facial expressions, and 5 types of brain waves) of learners are measured during visual- and text-based programming language learning. According to the experimental results, the values of "sadness" and "brain wave representing difficulty" were higher when using a visual-based programming language than when using a text-based programming language. Furthermore, the difference was larger in the group with poor keyboard input skills. This group performed the task while feeling contempt, sadness, negative emotions, and difficulties.**

***Index Terms*—e-Learning, Language Learning, Learning Condition, Heart Rate, Facial Expression Analysis, EEG**

I. Introduction

In recent years, visual-based programming languages (hereafter VPL) have been used as an introduction to programming. Thereafter, there is a transition to text-based programming languages such as C and Java (hereafter TPL). However, a seamless transition method has not been established.

We have started a research project aimed at establishing a methodology for the transition from VPL to TPL. Specifically, in this project, we will investigate and prototype educational content that combines the benefits of learning VPL and TPL and bridges the gap between the two (referred to as an IPL), and through demonstration experiments, we will assess not only the learning outcomes but also the learning state during the learning process. This will allow us to evaluate the effectiveness of the IPL and complete the educational content (IPL) useful for future primary and secondary programming education. With regard to the evaluation of IPL, which is supposed to bridge the gap between VPL and TPL, previous studies have focused only on the results of learning effects, such as post-learning questionnaires and grades, and have only assessed whether the students were able to understand the language. However, these evaluation methods cannot accurately measure the effects of IPL. In our research project, in addition to conventional evaluation methods after learning, we will measure the learning state using biometric information such as electroencephalography (EEG), eye gaze, and facial expressions during learning, as well as learning history during learning. We will then analyze and assess whether IPL plays an intermediate role between VPL and TPL and contributes to a smooth transition. Once established, this research is expected to enable beginners of programming languages to start learning a VPL and then seamlessly and spontaneously transition to learning a TPL.

In our previous study [4], we first proposed intermediate content as part of the research project described above. Then, through empirical experiments, we demonstrated that learning with our proposed intermediate content between learning VPL and TPL leads to an improved comprehension of the TPL afterward. Furthermore, the proposed intermediate content was evaluated using a questionnaire for its characteristics intermediate between VPL and TPL. In our previous study, we measured learners' biometric data (EEG, heart rate (HR), and facial expressions) while learning VPL and TPL and investigated whether there were any differences in the biometric data while learning the two languages. Multiple regression analysis using EEG as the objective variable and HR and 10 facial expressions as explanatory variables revealed differences in the positive and negative signs of the coefficients of the explanatory variables.

In this study, following our previous study [6], we measure 18 types of biometric information (HR, 12 types of facial expressions, and 5 types of brain waves (β/α)) to analyze the differences in biometric information while learning VPL and TPL.

II. Previous work

Xu et al. investigated existing academic databases and observed the overall impact of block-type VPL and TPL environments on both cognitive (achievement, problem-solving, etc.) and affective (satisfaction, confidence, motivation, etc.) learning outcomes of students. However, they could not show the statistical advantages of block-type VPL use and efficiency for novice programmers, but they stated the importance of further research into hybrid languages [11].

Tóth et al. highlighted the existence of a gap between VPL and TPL. They observed the migration from a VPL (MIT App Inventor 2) to a TPL (Android Studio) using Java Bridge Code Generator as a mediator of knowledge transfer. They claimed that the gap between VPL and TPL was bridged by the Java Bridge Code Generator [3].

Weintrop et al. experimentally assessed changes in knowledge transfer between learners who started with VPL and those who started with TPL and observed that there was no significant difference between the two types of language. The study was not about the transition from VPL to TPL, but a comparative study of knowledge after mastering TPL skills [10].

In our previous study [4], we demonstrated, through empirical experiments, that learning a TPL is better when our proposed intermediate content is inserted between the learning of a VPL and the learning of a TPL. Furthermore, a questionnaire was used to assess the characteristics of the proposed intermediate content, which is positioned between VPL and TPL.

In addition, we are promoting research on how to use EEG to monitor the learning progress of learners. We have obtained EEG information during the performance of a keyboard typing task and have demonstrated that the value of β/α increases when the task is difficult [8] [9]. In this study, considering the evaluation of EEG, we confirmed that there is a difference in EEG when solving problems in a VPL (Scratch) and a TPL (C). Specifically, for the VPL, the value of β/α did not increase as the task became increasingly difficult. This result made us realize that different pathways of thought may be used during the learning processes of VPL and TPL [5].

III. Experimental methods

A. Experimental participant

The participants in this experiment were seven fourth-year students at Shonan Institute of Technology. They had studied programming-related courses in the same department at the same university for several years. Their programming skills are nearly identical. However, the analysis revealed that the students' keyboarding skills varied.

B. Web services used in the experiment

For programming the VPL, we used Google blockly [1] (see Figure 1). This site contains puzzles, mazes, and other tasks, but to match the content of the TPL, we decided to work on the music task. In addition, we used JSFiddle [2] (see Figure 2) for programming the TPL. This site is an integrated environment for executing the JavaScript language, a TPL. By adding the Beepplay library as a resource setting, music with beep sounds can be created.

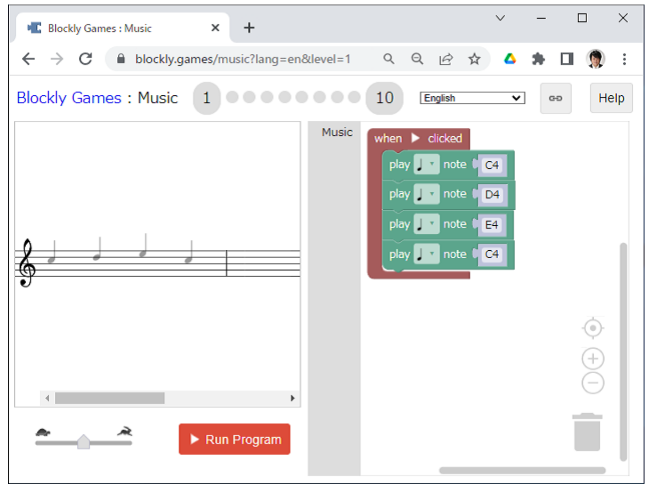

Fig. 1. Screen of Google blockly (Misic)

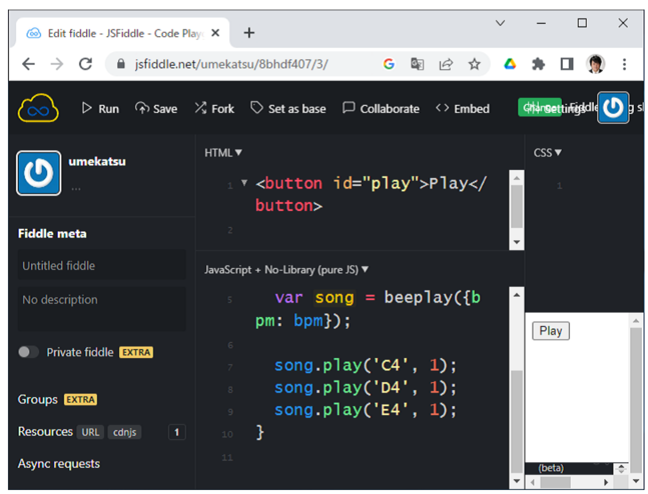

Fig. 2. Screen of JSFiddle (Beeplay)

C. Tasks to be used in the experiment

First, one practice session was performed, followed by two main experiments. Specifically, the task was to program songs so that they sounded the same as in the score. The songs used in each exercise and experiment are listed in Table I. These scores were printed and presented to the participants before the experiment. The scores used in the experiments are written in international notation (C3, C4, etc.) above the notes. Figures 3-5 show a part of the score of each song.

TABLE I
SONGS USED IN THE EXPERIMENT

Experiment	Song title
Practise	Froggy's Song
Experiment 1	Mary Had A Little Lamb
Experiment 2	Jingle Bells

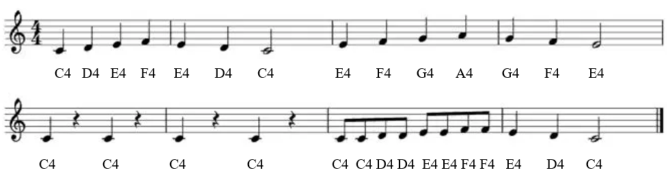

Fig. 3. Score of Froggy's Song

Fig. 4. Score of Mary Had A Little Lamb

Fig. 5. Score of Jingle Bells

D. Biological equipment used in the experiment

1) Simple EEG: NeuroSky's MindWave Mobile 2 is used to measure brain waves. EEG data are collected through a Bluetooth connection between the headset and the ThinkGear Connector, and the logging application communicates with the ThinkGear Connector over TCP/IP. The ThinkGear Connector is a driver that provides communication with NeuroSky's MindWave Mobile 2. The EEG data acquired using the EEG monitor are δ waves (0.5–2.75 Hz), θ waves (3.5–6.75 Hz), low α waves (7.5–9.25 Hz), high α waves (10–11.75 Hz), low $beta$ waves (13–16.75 Hz), highβwave(18–29.75 Hz), low γ wave (31–39.75 Hz), and medium γ wave (41–49.75 Hz). Each value is a unitless 4-byte floating decimal number. In this study, we use four types of waves, namely, low α waves, high α waves, low β waves, and high $beta$ waves, in accordance with the previous study [12]. Hereafter, low α waves are denoted as α_l, high α waves as α_h, low β waves as β_l, and high β waves as β_h. In this study, we use five types of β/α values from the existing study [7], including four combinations of β/α (β_l/α_l, β_h/α_l, β_l/α_h, β_h/α_h) and $(\beta_l + \beta_h)/(\alpha_l + \alpha_h)$ as the average of low and high frequencies.

2) Heart rate monitor: Garmin's "Venu 2" is used to measure HR. This device is a wristwatch-type device that can measure HR, respiration rate, stress, and so on. Its data can be exported by linking it to a dedicated smartphone application. Here, we use HR among the measured values.

3) Facial recognition: To determine facial expressions, it is necessary to record and analyze facial expressions. A Logitech C920n webcam is used to record facial expressions. For facial expression analysis, CAC's Kokoro sensor (equipped with Affectiva's emotion recognition engine) is used to analyze the captured video. This application recognizes human faces from captured or real-time videos and can quantify 12 types of emotions (anger, contempt, disgust, fear, joy, sadness, surprise, sentimentality, confusion, neutral, engagement, and valence (positive and negative expressions)). These emotions except valence are expressed as a number between 0 and 100, whereas valence is expressed as a number between -100 and 100.

4) Keylogger: The mouse is used more frequently for the VPL, and the keyboard is used more frequently for the TPL. To observe the frequency of use of each input device, a keylogger was installed in the experimental PC, and the use of each input device was measured during the experiment.

E. Experimental flow

Figure 6 depicts the overall flow of the experiment. As depicted in Figure 6, the participants were randomly divided into two groups, one working from TPL (Group A) and the other from VPL (Group B). During each task, HR, facial expression recording (facial expression analysis is done after the experiment), and EEG are measured. For HR and EEG, one data point is acquired approximately every second. For facial expression analysis, 24 data points are acquired per second as a result of analyzing the video. Although omitted from Figure 6, HR variability (HRV) was measured for 30 s each before and after a 5-min rest period. HRV was not included in the analysis of this study.

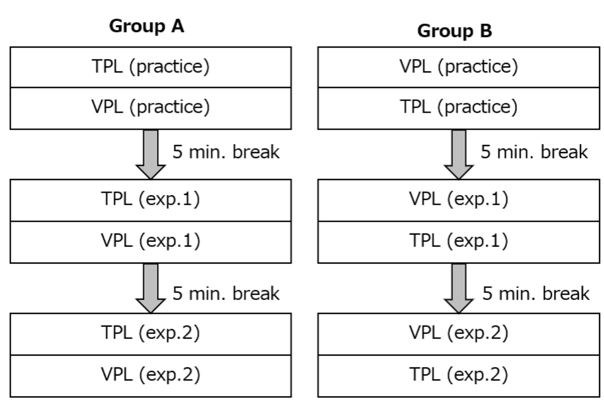

Fig. 6. Experimental flow

IV. Experimental results

Seven participants, four from Group A (participants A–D) and three from Group B (participants E–G), solved four questions: each participant solved exp.1 in TPL, exp.1 in VPL, exp.2 in TPL, and exp.2 in VPL. For each of the four experiments, data on HR, 12 emotions, and 5 brain waves were obtained. Table II lists a part of the experimental results for Participant A's HR.

TABLE II
A part of the experimental results of the HR of Participant A

Experiment	Raw data	Z-score
TPL (exp.1)	70	0.9581
TPL (exp.1)	72	1.6484
⋮	⋮	⋮
VPL (exp.1)	65	−0.7677
VPL (exp.1)	66	−0.4225
⋮	⋮	⋮
TPL (exp.2)	68	0.2678
TPL (exp.2)	67	−0.0774
⋮	⋮	⋮
VPL (exp.2)	77	3.3742
VPL (exp.2)	78	3.7194
⋮	⋮	⋮
Mean value	67.2242	0.0000
Standard deviation	2.8972	1.0000

The magnitude of the biometric values varies depending on the type of biometric data and the differences among the participants. Therefore, the analysis in the subsequent sections is based on the values converted from the raw biometric data to Z-scores. Specifically, as listed in Table II, the mean value and standard deviation are obtained for each participant in the experiment and each type of biometric information and converted to a Z-score.

V. Analysis

A. Confirmation that there is no difference between Groups A and B

We want to confirm that there is no difference in the experimental results between Group A, which works with TPL, and Group B, which works with VPL. For all biometric information, we tested whether there was a difference between the mean values of the TPL of Group A (four persons $\times$ 2 tasks $=$ 8 data) and Group B (three persons $\times$ 2 tasks $=$ 6 data). Similarly, we tested whether there was a difference between the mean of the VPL of Group A (eight data points) and Group B (six data points). Wilcoxon's rank sum test was used for the test. The results showed that among HR, 12 emotions, and 5 types of EEG, only 3 types of EEG, β_h/α_l, β_h/α_h, and β_{l+h}/α_{l+h}, were significantly higher in Group A when performing tasks with VPL. However, the other 15 types of biometric data were not significantly different between the TPL and VPL performance. Therefore, in the analysis that follows, Groups A and B are considered identical without distinguishing between them.

B. Analysis of differences in keyboarding experience

In this experiment, keyboard and mouse inputs were monitored by a keylogger during each task. The logs showed that Participants A, E, F, and G used Ctrl-C (copy) and Ctrl-V (paste) frequently during the TPL task. Conversely, Participants B, C, and D did not use these shoot-cut keys at all. We hypothesized that this experience with keyboard input may have influenced the changes in the biometric data. Therefore, in the next section, we analyze the results by dividing the group good at keyboard input (AEFG) and the group poor at keyboard input (BCD).

C. Differential analysis of TPL and VPL

1) Difference analysis for all: Before analyzing the data for Groups AEFG and BCD, we analyzed the data for all participants. Using the experimental data for all participants (7 participants, 14 data points in total), we test whether there is a difference between the biometric information obtained when solving exp.1 and exp.2 in TPL and when solving exp.1 and exp.2 in VPL. The test method is the Wilcoxon signed rank test on paired two-sample data, which is a nonparametric method that does not assume a specific distribution such as a normal distribution for the population distribution. The results of the test are presented in Table III.

TABLE III
Differential analysis of TPL and VPL for all participants

Biometric data	Mean value		Test statistic U	Two-side p-value
	TPL	VPL		
HR	0.0514	−0.0227	56	0.8261
anger	−0.0630	0.0553	35	0.2719
contempt	−0.0894	0.0927	13	0.0132*
disgust	−0.057	0.0478	35	0.2719
fear	−0.0735	0.0617	36	0.3003
joy	0.0378	−0.0421	59	0.6832
sadness	−0.0958	0.0939	18	0.0303*
surprise	0.0067	−0.0084	70	0.2719
sentimentality	0.0116	−0.0164	58	0.7299
confusion	0.0635	−0.0628	83	0.0555
neutral	−0.0053	0.0126	55	0.8753
engagement	−0.0263	0.0120	45	0.6378
valence	0.1246	−0.0574	79	0.0962
β_l/α_l	0.1709	0.2279	28	0.1240
β_l/α_h	0.2351	0.1702	82	0.0640
β_h/α_l	0.1109	0.3069	18	0.0303*
β_h/α_h	0.1623	0.2247	46	0.6832
β_{l+h}/α_{l+h}	0.1250	0.1775	37	0.3305

The values for "contempt," "sadness," and "β_h/α_l" were significantly higher when the task was solved in a VPL.

2) Differential analysis to account for differences in keyboarding experience: As mentioned in the previous section, the biometric information may change depending on whether the user is good or poor at keyboard input. In this section, we conduct Wilcoxon signed rank tests on two-sample data with the same correspondence as in the analysis presented in section V-C1, separately for Groups AEFG and BCD. For Group AEFG (good at keyboard input), there was no significant difference between TPL and VPL for all biometric information. The test results for Group BCD (the group with poor keyboard input) are presented in Table IV.

TABLE IV
DIFFERENTIAL ANALYSIS OF TPL AND VPL FOR GROUPS THAT HAVE DIFFICULTY WITH KEYBOARD INPUT

Biometric data	Mean value TPL	Mean value VPL	Test statistic U	Two-side p-value
HR	0.2507	−0.1828	15	0.3454
anger	−0.0710	0.0707	3	0.1159
contempt	−0.1273	0.1398	0	0.0277*
disgust	0.0113	−0.0175	14	0.4631
fear	−0.0980	0.0908	4	0.1730
joy	0.0401	−0.0538	12	0.7532
sadness	−0.1905	0.1886	0	0.0277*
surprise	−0.0579	0.0467	10	0.9165
sentimentality	0.0669	−0.0829	16	0.2489
confusion	0.0678	−0.0683	18	0.1159
neutral	−0.0373	0.0529	8	0.6002
engagement	0.0187	−0.0384	11	0.9165
valence	0.2636	−0.1111	21	0.0277*
β_l/α_l	0.4222	0.5091	5	0.2489
β_l/α_h	0.5302	0.4137	21	0.0277*
β_h/α_l	0.2942	0.6815	0	0.0277*
β_h/α_h	0.3307	0.5722	0	0.0277*
β_{l+h}/α_{l+h}	0.2536	0.4532	0	0.0277*

As shown in Table IV, the BCD group (the group with poor keyboarding skills) exhibited significantly higher values for “contempt” and “sadness” when performing the task in the VPL than when using the TPL. Furthermore, “valence (positive and negative emotions)” were significantly lower. In other words, the AEFG group (good at keyboard typing) exhibited no significant difference between TPL and VPL. However, the BCD group (the group with poor keyboarding skills) performed the task with negative emotions, such as contempt and sadness, when using VPL. In addition, three types of EEG, β_h/α_l, β_h/α_h, and β_{l+h}/α_{l+h}, were significantly higher when the task was solved in a VPL. In other words, the participants felt more difficulty when using a VPL.

D. Differential analysis of groups good at keyboarding and groups poor at keyboarding

1) Difference between TPL and VPL: In the previous sections, we analyzed the differences in biometric information when using TPL and VPL. In this section, we analyze the differences between Groups AEFG and BCD. In this analysis, we do not treat the TPL and VPL as separate data, but rather treat the difference between the TPL and VPL. As an example, Table V shows the difference between the TPL and VPL for HR. In the following analysis, the “difference” shown in the rightmost column of Table V is used.

2) Differential analysis of group good at keyboarding and group poor at keyboarding: Wilcoxon’s rank sum test is used to check whether there is a difference between the mean difference between the TPL and VPL of the four participants in the AEFG group and the three participants in the BC group. The results of the test are presented in Table VI.

TABLE V
MEAN DIFFERENCE BETWEEN TPL AND VPL OF HR

AEFG	Exp.	TPL	VPL	difference
Participant A	exp.1	−0.0069	−0.7849	0.7780
Participant A	exp.2	0.3125	0.4788	−0.1663
Participant E	exp.1	−0.2894	0.3107	−0.6001
Participant E	exp.2	−0.6287	0.6162	−1.2449
Participant F	exp.1	0.2135	0.3413	−0.1278
Participant F	exp.2	−0.9341	0.3789	−1.3130
Participant G	exp.1	0.9242	0.0853	0.8389
Participant G	exp.2	−0.3755	−0.6475	0.2720
Mean value of 8 data		−0.0981	0.0973	−0.1954

BCD	Exp.	TPL	VPL	difference
Participant B	exp.1	0.9833	−0.8194	1.8027
Participant B	exp.2	0.5128	−0.2585	0.7713
Participant C	exp.1	−0.6591	−0.0796	−0.5795
Participant C	exp.2	0.1993	0.5215	−0.3222
Participant D	exp.1	0.0317	0.0353	−0.0037
Participant D	exp.2	0.4361	−0.4960	0.9321
Mean value of 6 data		0.2507	−0.1828	0.4335

TABLE VI
DIFFERENTIAL ANALYSIS OF GROUP GOOD AT KEYBOARDING AND GROUP POOR AT KEYBOARDING

Biometric data	Mean of difference AEFG	Mean of difference BCD	Test statistic U	Test statistic Z	Two-side p-value
HR	−0.1954	0.4335	15	1.1619	0.2453
anger	−0.1007	−0.1417	19	0.6455	0.5186
contempt	−0.1184	−0.267	12	1.5492	0.1213
disgust	−0.2051	0.0289	12	1.5492	0.1213
fear fear	−0.0950	−0.1889	19	0.6455	0.5186
joy	0.0696	0.0939	19	0.6455	0.5186
sadness	−0.0476	−0.3792	3	2.7111	0.0067**
surprise	0.1049	−0.1046	15	1.1619	0.2453
sentimentality	−0.0634	0.1499	13	1.4201	0.1556
confusion	0.1190	0.1361	23	0.1291	0.8973
neutral	0.0364	−0.0902	19	0.6455	0.5186
engagement	−0.1098	0.0570	19	0.6455	0.5186
valence	0.0375	0.3747	9	1.9365	0.0528
β_l/α_l	−0.0346	−0.0869	19	0.6455	0.5186
β_l/α_h	0.0261	0.1165	9	1.9365	0.0528
β_h/α_l	−0.0527	−0.3872	8	2.0656	0.0389*
β_h/α_h	0.0719	−0.2415	9	1.9365	0.0528
β_{l+h}/α_{l+h}	0.0579	−0.1997	9	1.9365	0.0528

As shown in Table VI, there is a significant difference in “sadness” and “β_h/α_l.” The mean difference between Groups AEFG and BCD was negative, but the absolute value of the difference was larger in Group BCD. In other words, the BCD group (the group with poor keyboarding skills) exhibited significantly greater sadness and β_h/α_l when learning the VPL.

VI. CONSIDERATION

The table of differential analysis of TPL and VPL for all participants (Table III) shows that the values of “sadness” and “β_h/α_l” are higher for VPL than for TPL. The difference was larger for the BCD group than for the AEFG group from Table VI.

From the differential analysis of TPL and VPL for groups with difficulty with keyboard input (Table IV), the BCD group showed significantly higher "contempt" and "sadness" and significantly lower "valence." The results showed that the participants performed the task with negative emotions, such as contempt and sadness. The three types of β/α values were also significantly higher in EEG, indicating that the participants felt more difficulty when using the VPL.

VII. Conclusions and future work

In this study, we measured learners' HR, 12 facial expressions, and 5 brain waves (β/α) to determine whether there were any differences in biometric information while learning both VPL and TPL while performing tasks. We found that the values of "sadness" and "difficulty EEG (β_h/α_l)" were higher when learners used VPL than TPL. The difference was larger in the group that had difficulty with keyboard input. Furthermore, the group that had difficulty with keyboard input performed the task with contempt, sadness, and negative emotions, as well as with a sense of difficulty.

In the future, conducting experiments with more participants will be necessary. In addition, this experiment was conducted with fourth-year university students who had experience with TPL. Therefore, future experiments should be conducted with junior and senior high school students who are in the transition period from a VPL to a TPL. The conclusion of this study demonstrates that for developing an intermediate language that can bridge the gap between a VPL and a TPL in the future, it is necessary to develop an intermediate language that does not create stress and negative emotions for students at various levels of understanding and proficiency. We believe that in the future we must clarify how to improve our teaching and how to generalize the results of this study.

Research ethics

All experiments were approved by the Research Ethics Committee of Shonan Institute of Technology. We received written informed consent from the participants and their parents or guardians.

Acknowledgments

Part of the work reported here was conducted as a part of the research project "Research on e-learning for next-generation" of Waseda Research Institute for Science and Engineering, Waseda University. Part of this work was supported by JSPS KAKENHI Grant Numbers JP22H01055, JP21K18535, and JP20K03082. Research leading to this paper was partially supported by the grant of "ICT and Education" of JASMIN.

References

[1] Blockly games. https://blockly.games/. Accessed: 17 January 2023.

[2] Jsfiddle. https://jsfiddle.net/. Accessed: 17 January 2023.

[3] T. Tóth and G. Lovászová. Mediation of knowledge transfer in the transition from visual to textual programming. *Informatics in Education*, 20:489–511, 2021.

[4] K. Umezawa, K. Ishida, M. Nakazawa, and S. Hirasawa. A proposal and evaluation of intermediate content for transition from visual to text-based languagesg. *The Institute of Electronics, Information and Communication Engineers (IEICE) Technical report*, pp. 1–7, 2022.

[5] K. Umezawa, Y. Ishii, M. Nakazawa, M. Nakano, M. Kobayashi, and S. Hirasawa. Comparison experiment of learning state between visual programming language and text programming language. *2021 IEEE International Conference on Engineering, Technology & Education (TALE)*, pp. 01–05, 2021.

[6] K. Umezawa, M. Nakazawa, and S. Hirasawa. About the comparison of biometric information during learning of visual- and text-based programming language. *The Institute of Electronics, Information and Communication Engineers (IEICE) Technical report*, pp. 1–6, 2023.

[7] K. Umezawa, T. Saito, T. Ishida, M. Nakazawa, and S. Hirasawa. An electroencephalograph-based method for judging the difficulty of a task given to a learner. *Proceeding of the IEEE 17th International Conference on Advanced Learning Technologies (IEEE ICALT 2017)*, pp. 384–386, 2017.

[8] K. Umezawa, T. Saito, T. Ishida, M. Nakazawa, and S. Hirasawa. Learning state estimation method by browsing history and brain waves during programming language learning. *Trends and Advances in Information Systems and Technologies. WorldCIST'18 2018. Advances in Intelligent Systems and Computing, Springer, Cham.*, 746, 2018.

[9] K. Umezawa, T. Saito, T. Ishida, M. Nakazawa, and S. Hirasawa. Learning-state-estimation method using browsing history and electroencephalogram in e-learning of programming language and its evaluation. *Proceeding of the International Workshop on Higher Education Learning Methodologies and Technologies Online (HELMeTO 2020)*, pp. 22–25, 2020.

[10] D. Weintrop and UriWilensky. Transitioning from introductory block-based and text-based environments to professional programming languages in high school computer science classrooms. *Computers & Education*, 142, 2019.

[11] Z. Xu, A. D. Ritzhaupt, F. Tian, and K. Umapathy. Block-based versus text-based programming environments on novice student learning outcomes: A meta-analysis study. *Computer Science Education*, 29:177–204, 2001.

[12] K. Yoshida, Y. Sakamoto, I. Miyaji, and K. Yamada. Analysis comparison of brain waves at the learning status by simple electroencephalography. *InAdvances in Knowledge-Based and Intelligent Information and Engineering Systems 2012*, pp. 1817–1826, 2012.