Experimentally Analyzing Relationships between Learner's Status in the Skill Acquisition Process and Physiological Indices

Yoshimasa Ohmoto
Department of Intelligence Science and Technology
Graduate School of Informatics
Kyoto University
Kyoto, Japan
Email: ohmoto@i.kyoto-u.ac.jp

Takahiro Matsuda
Department of Intelligence Science and Technology
Graduate School of Informatics
Kyoto University
Kyoto, Japan
Email: matsuda@ii.ist.i.kyoto-u.ac.jp

Toyoaki Nishida
Department of Intelligence Science and Technology
Graduate School of Informatics
Kyoto University
Kyoto, Japan
Email: nishida@i.kyoto-u.ac.jp

***Abstract*—There are many tasks for which people need domain-specific skills learned through long-term practice. Many skill acquisition models have been proposed that include the mental states of the learners, but only few studies have tried to estimate mental states using objectively measured data. The purpose of this study was to experimentally investigate whether the status of the skill acquisition process could be estimated by physiological indices that indicate the learner's mental states. For this purpose, we conducted an experiment to obtain data on physiological indices and a subjective report of the feeling of difficulty during the skill acquisition task. As a result, we confirmed the relationship between the participant's statuses and the physiological indices. In addition, we classified the trials with the feeling of difficulty in the experiment using Support Vector Machine. As a result, the values of accuracy were over 0.65 when the data not used to calculate the SVM model were classified. We could show that the physiological indices are helpful clues to estimate the status of the skill acquisition process.**

Keywords–Skill acquisition model; Mental state estimation; Physiological indices.

I. INTRODUCTION

This article is a substantially extended version of the authors' paper "Experimentally Analyzing the Skill Acquisition Model using Task Performance and Physiological Indices" [1], presented at COGNITIVE 2017, the Ninth International Conference on Advanced Cognitive Technologies and Applications, IARIA in Greece.

A skill is the ability to perform a task with pre-determined results, often involving a given amount of time, energy, or both. There are many tasks for which people need domain-specific skills learned through long-term practice. We call this kind of undertaking a "skill task." To acquire the abilities for a skill task, people usually train by carrying it out repeatedly. However, it is difficult to learn the aptitude in question alone because, in many cases, people cannot objectively monitor their skill level and task performance. Experts and instructors can support learners, but the methods to support learning a skill task have been established for few endeavors. Some systems have been proposed to support learning based on task performance.

Generally, when a person has acquired a skill, i.e., proficiency or expertise, this means that person can efficiently carry out a specific task. Previous studies, e.g., [2], are founded on the principle that learners pass through five stages: Novice, competence, proficiency, expertise, and mastery. The phases are characterized by how rules interplay with real-world context. A novice will simply follow the rules that they are given and not consider context. Intermediate stages contain a mix of rule following, combined with more and more sophisticated consideration of context [3]. To assess skill level, the skill task is divided into a number of sub-tasks, the performance on which is rated. However, to competently perform a job, many skill tasks require various sub-skills corresponding to the sub-tasks. In this case, even when the learner has acquired some sub-skills, a situation may arise where he/she cannot synthetically use them. In addition, it is often difficult to rate skill task performance itself. Tranquillo and Stecker [3] pointed out that a master may, in fact, not be able to state the rules or the heuristics that they are using. For example, we cannot rate presentation skills independent of the contents of the presentation. To support the skill acquisition of the skill task, identifying the statuses in the skill acquisition process, e.g., the steep acceleration section or the plateau section in a learning curve, is more useful than rating the skill level. To identify this status, human instructors focus on the learner's responses to unknown situations (such as questions and answers), giving a new challenge, and deliberately making mistakes. In other words, they evaluate the skill task model that the learner has. Of course, the skill task model cannot be observed directly.

The important point is that we can use the learner's own recognition of his/her skill task model. The Dreyfus model of skill acquisition [2] shows how students acquire abilities through formal instruction in addition to practicing. This model identifies skill level based on four binary qualities: (1) recollection (non-situational or situational); (2) recognition (decomposed or holistic); (3) decisions (analytical or intuitive); and (4) awareness (monitoring or being absorbed). The model is intuitive, but no one knows how to evaluate the four factors concretely. We considered an approach to estimate the learner's recognition, i.e., the learner's mental state, from objectively measured data during the skill acquisition process. For this assessment, we used physiological indices that could gauge the responses of the learner's autonomic nervous system related to his/her mental states. The physiological indices are usually employed to ascertain mental stress.

Kraiger et al. [4] attempted to move toward a training evaluation model by developing a classification scheme for evaluating learning outcomes. They integrated theory and research from a number of diverse disciplines and provided a multidimensional perspective on learning outcomes. They provided a classification scheme for learning outcomes for training evaluation but were unable to propose a method to identify the outcomes objectively. Mitchell et al. [5] assessed participants' self-efficacy goals, expected performance, and the degree to which certain judgments required cognitive processing. The results showed that self-efficacy was a better predictor of performance than were expected score or goals on early trials, whereas the reverse was true for later trials. This result showed that the mental state is useful for estimating the skill level because the responses changed with the skill acquisition. From these results, we think that we needed a method to objectively estimate the learner's mental state that changed with the skill acquisition process.

The final goal of this study was to develop a method to determine the status of the skill acquisition process in detail for the skill task, using task performance and the learner's mental states. For this purpose, we experimentally investigated whether the status of the skill acquisition process could be estimated by the physiological indices that can be used to estimate the learner's mental states. We conducted an experiment to obtain the data of the physiological indices and a subjective report of the learner's mental states.

The paper is organized as follows. Section II briefly introduces previous works on the skill acquisition model and assessing human stress. Section III explains the outline of the technique for appraising human stress and the skill acquisition model via two dimensions: (1) task performance and (2) the learner's mental states. Section IV describes the experiment and the analysis of the data. Section V discusses the analysis, and Section VI lays out our conclusions and future works.

II. RELATED WORK

The Dreyfus Model of skill acquisition [2] seems to be one of the most famous models of how students acquire skills through formal instruction and practicing. This model is used in a broad range of tasks, such as defining an appropriate level of competence, supporting judgements of when a learner is ready to teach others, and so on. However, there are some criticisms of this model [6]. According to these authors, there is no empirical evidence for the presence of stages in the development of expertise.

Day et al. [7] examined the viability of knowledge structures as an operationalization of learning in the context of a task that required a high degree of skill using a complex video game. At the end of acquisition, participants' knowledge structures were assessed and the similarity of trainees' knowledge structures to an expert structure was correlated with skill acquisition and was predictive of skill retention and skill transfer. In addition, knowledge structures mediated the relationship between general cognitive ability and skill-based performance. However, they did not mention the process of change of the cognitive ability during the skill acquisition.

Langan-Fox et al. [8] summarized traditional models ([9][10][11]). They pointed out that models of skill acquisition largely ignore the experiences and dynamic internal processes of a person while learning a skill. They attempted to highlight the importance of a dynamic description of skill acquisition in their research. Process-oriented factors such as motivation, memory, interruptions, emotion, and metacognition are investigated in relation to skilled performance. However, their discussions were conceptual and they did not investigate the matter experimentally.

There are also some studies focusing on the learning process, e.g., [12][13][14]. This is called the learning curve, and it is well known that skills are acquired gradually, followed by a period of steep acceleration and then a plateau phase. However, there is no definite answer as to why such stepwise learning occurs. In general, it is thought that in the learning status where changes in cognition for tasks are required, learning is in the plateau section until the change occurs. However, it is difficult to objectively observe how the change occurs. Observing changes in mental state during the skill acquisition is also important in clarifying such processes.

There are some studies in which the physiological indices were used for identifying the mental states. Patterson et al. [15] used heart rate variability (low frequency:high frequency ratio) to differentiate invested cognitive effort during the acquisition and retention of a novel task. They found the usefulness of heart rate variability in discriminating the cognitive effort invested for a recently acquired skill. Walker et al. [16] sought to identify a physiological measure that could help predict team performance during a complex and dynamic task. Regression analyses showed that team autonomic activity accounted for 10% of the variance in team performance scores. They showed that the task performance can be predicted from physiological indices. Parsinejad [17] tried to infer the mental workload changes of human operator using physiological measurements and performance metrics, while keeping participants' inexperience as the key parameter. This research showed that physiological indices and the participant's finger-stroke patterns on the touch screen could help flag unfamiliarity of participants in the difficult game. These studies showed that the physiological indices were useful to identifying the learner's mental states and they could predict the task performance based on the physiological indices and the learner's behavior. However, they did not focused the relationships between the mental states and the learning process.

III. THE SKILL ACQUISITION MODEL, INCLUDING MENTAL STATES

Some previous studies have proposed the skill acquisition model, which considers learners' mental states [2][8]. However, the mental states were evaluated by human observation. In other words, previous studies have not focused on how to measure, evaluate, and use mental states via objective approaches to the skill acquisition model. In addition, they have centered on mental states through the lens of a specific skill level, but have not concentrated on the dynamics of people going through a learning process. This study aimed to develop a technique for assessing a learner's skill level based on his/her performance and mental states. To do so, we first confirmed the relationship between subjective reports on mental states and the measured physiological indices. We then interpreted the data and performance for estimating skill levels and the process of learning. In this section, we explain the physiological indices that were used to appraise mental states, and propose a skill acquisition model that has two dimensions that define skill level.

A. Physiological indices for assessing mental states

In this study, we propose the skill acquisition model using task performance and mental states. We especially focus on how to gauge and use mental states in the skill acquisition model. It is hard to use a learner's behavior to evaluate mental states because the learner's behavior depends on the task and skill level. However, when a learner feels that an activity is difficult, he/she feels mental stress due to his/her line of thinking and stimuli from the endeavor. Therefore, we consider stress useful for appraising a typical mental state in the learning process.

Many previous investigations have reported on physiological indices for estimating mental stress. However, in ongoing daily interactions, we can often find concurrent physiological responses that are not related to the event. One reason is that people involved in continuous interactions often plan their actions, such as what to tell and how to move. Therefore, to assess human mental states, we had to consider the context of an interaction and the response characteristics of the physiological indices.

Physiological indices are biological reactions caused by the autonomic nervous system, for example: brain waves, potential differences in cardiographs, variations in blood pressure, pulse waves, respiration, body temperature, muscle potential, and skin conductance. In continuous interactions, some of these are susceptible to noise from body motions. We used skin conductance responses (SCR) and electrocardiograms (Low Frequency (LF)/High Frequency (HF) values) because these are relatively resistant to noise.

Since the underlying mechanisms of SCR and electrocardiograms are different, we expected that they could be used to distinguish between different responses from various sources of stress. Sweating is controlled by the sympathetic nervous system [18] and can be elicited by emotional stimuli, intellectual strain, or painful cutaneous stimulation. The underlying mechanisms of SCRs are more related to anticipation, expectation, and attention concentration [19]. We thus anticipated that SCRs could be used to tell when someone is dealing with an unexpected situation.

For electrocardiograms, the LF/HF value is calculated using instantaneous heart rate. It shows heart rate variability (HRV), which is controlled by the sympathetic and parasympathetic nervous systems and humoral factors. The underlying mechanisms of HRV are complex. Lacey and Lacey [20] suggested that it is caused by sensory intake and sensory rejection. In addition, the parasympathetic nervous system responds quickly (< 1 s) to stimuli. We thus thought that the LF/HF (HRV) would show reactive responses based on external stimuli.

B. Relationship between a mental state and skill acquisition

The performance on a skill task is often difficult to rate. We consider skill acquisition to mean that a learner has constructed an efficient task model. If we could directly evaluate the learner's skill task model, we could identify the statuses of the skill acquisition process. Human instructors focus on the learner's responses to unknown situations to estimate the skill task model. The learner's responses are expressions that the learner feels difficulty and senses the patterns of mistakes. Since learners might recognize the feeling of difficulty and the mistakes consciously or unconsciously, we might estimate them based on the learner's mental states.

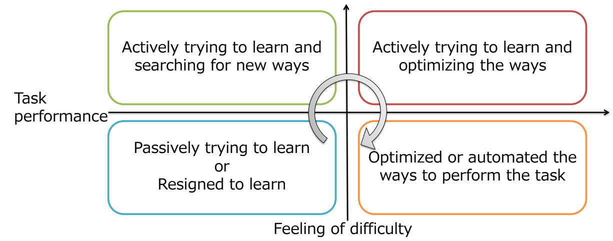


Figure 1. The outline of the skill acquisition model via two dimensions.

We focused on mental state changes to estimate whether the learners thought their constructed task models were efficient. In general skill acquisition processes, we can draw a learning curve whose shape resembles a series of steps, e.g., [12][13][14]. In the steep acceleration section of the curve, the learners receive feedback from subjective and objective evaluations and improve their behavior. In the plateau section, the learners optimize their behavior and sometimes execute trial and error for the next step to improve the task performance. When the learners are constructing the skill task model, they experience intrinsic stresses. When the learners are improving their model, they mainly experience extrinsic stresses. As mentioned above, we expect that the intrinsic and extrinsic stress can be estimated with physiological indices. By using these estimates, we may estimate whether the learners thought their constructed task models were efficient.

C. The skill acquisition model in two dimensions

We propose a skill acquisition model with two dimensions-task performance and mental state-when a person feels that a task is challenging (we call this a “feeling of difficulty”). Figure 1 shows the model and the two dimensions. In the traditional three phase model, Phase 1 corresponds to the lower left part, Phase 2 corresponds to the upper part, and Phase 3 corresponds to the lower right part. The “task performance” of the horizontal axis is the task result, objectively quantified. The “feeling of difficulty” of the vertical axis is a mental state, assessed based on the measured physiological indices. The traditional skill acquisition model assumes that task performance (weakly) rises with an increase in skill levels. When a skill can be segmented into small mutually independent sub-skills, this assumption may be true of the process of acquiring sub-skills. However, attaining sub-skills does not contribute to gaining more than a certain level of overall aptitude for many skill tasks, such as ballroom dancing. In this case, the synthetic use of sub-skills is often important, and indeed itself is a target of skill acquisition. When the learner practices synthetic use through trial and error, task performance often decreases. In our model, the trial and error process is included in the lower part. When the task performance decreases but the difficulty feeling is low, the model interprets the process as trial and error. We expected that the learner would cycle through this model throughout the skill acquisition process.

We assume that the skill acquisition process unfolds in the following manner. The initial state of the learner is in the upper or lower left part, i.e., performance is low. The learner usually needs trial and error to learn the skill task, so the state of the learner is maintained or transitions to the upper part. Through the learning process, task performance increases and

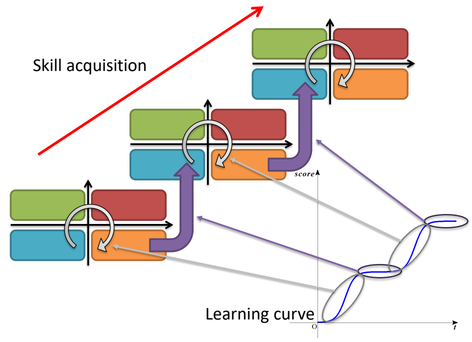
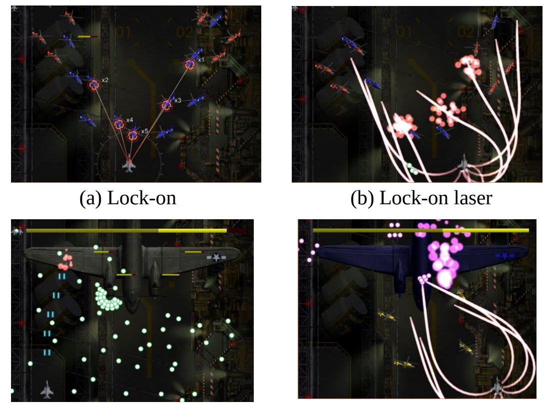


Figure 2. The outline of the skill acquisition process whose shape resembles a series of steps.



(a) Lock-on (b) Lock-on laser

(c) Gatling gun and enemy bullets (d) Final-boss battle

Figure 3. The screen shots of the shooting game.

the learner's state moves to the upper right part. In this state, the learner can perform the task at hand more efficiently than before, but he/she is not accustomed to or does not understand how to carry out the activity. Through practice at this stage, the learner can synthetically and automatically perform the task, and his/her state moves to the lower right part. In a common case, the learner continuously performs trial and error to find better ways, so task performance sometimes falls. In this case, the learner's state is ready for the next stage of the skill acquisition process. If the learner can find clues for better ways to perform the task, the skill acquisition process advances to the next phase. If the learner cannot find better ways, the skill acquisition process terminates in the lower right part of the stage. Figure 2 shows a series that demonstrates the flow of the skill acquisition process in our model. Through the cyclic skill acquisition process, learners deepen their understanding of the skills and the tasks. After that, they can find a breakthrough for better ways to perform the task, and advance to the next phase of the skill acquisition.

IV. Experiment

The purpose of this experiment was to investigate whether the physiological indices were related to the subjective reports about the feeling of difficulty toward the task, and whether we could evaluate the state of transition in the proposed skill acquisition model based on task performance and physiological indices. We adopted a shooter game as a skill task. Some previous studies, e.g., [7], have adopted various shooter games. The shooter game that we developed has features of the skill task. The advantages of using a shooter game as the skill task include the following: (1) The learner needs to obtain game playing skills, which is hard to verbalize; (2) We can control the difficulty of a task; and (3) We can easily analyze the skill acquisition process because we can independently divide the time series of the game events, which is the target of the skill acquisition. In addition, learners can repeatedly play the game with high motivation. We conducted an experiment in which the participants played the shooter game repeatedly, and we obtained the game scores and physiological indices during game play. After the experiment, we analyzed the data to confirm the relationship between the physiological indices and the subjective reports about the feeling of difficulty. Furthermore, we examined the relationship in the skill acquisition process.

A. *Task*

To achieve the best performance in the shooter game, the player must cultivate certain skills and gain certain knowledge, such as operation procedures, scoring rules, the way to defeat one's enemies, the features of game stages, basic survival patterns, and specific techniques to obtain a high score. Some of these cannot be verbalized and the best method varies among the players. To obtain game playing skills, the players must practice repeatedly.

Figure 3 shows the screen shots in the shooter game, which we developed. In this game, player uses two different methods of attack, Gatling gun (Figure 3 (c)) and Lock-on laser (Figure 3 (a)(b)). The Gatling gun is a quick and out-range attack method. When the player uses the Gatling gun to destroy the enemies, the game score is a minimum. The lock-on laser is powerful but needs a lock-on procedure near the enemies. When the player uses a lock-on laser to destroy enemies, the game score increases exponentially with the number of lock-on targets destroyed at the same time. The player tries to obtain as high a game score as possible by selectively using the two different attack methods. Of course, as the enemies are attacking the player, the lock-on laser cannot always be used to survive in the game.

In this experiment, the participants were only trained for the first stage of the shooter game. The stage was segmented into eight parts. The patterns of combinations and the movements of enemies were different in each part. The eighth part was the battle with the final boss enemy. There was a relaxation period after each part. The average clear time was designed to be about 150 seconds. The patterns are shown in Figure 4. The first part and the fourth part are relatively easy because the density of enemies is sparse and the moving patterns are simple. Since the densities of enemies in the second and third parts are thick and the moving patterns are complex, they are relatively difficult. From the fifth to the seventh parts, the densities of enemies are relatively sparse and the moving patterns are relatively simple, but special techniques are needed to get high scores. When the player used a suitable approach in any part, his/her score was several times higher than that obtained using an inappropriate procedure. When the enemies hit the player three times, the game was over. When the boss enemy was destroyed, the game was cleared. After the game

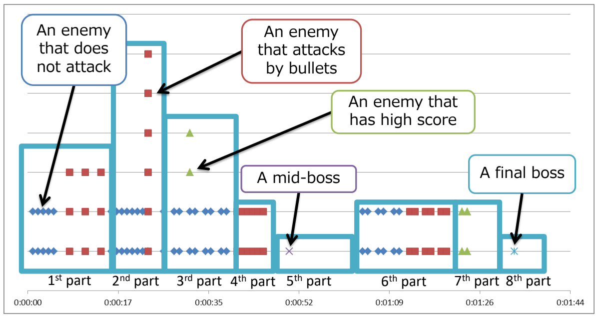


Figure 4. The outline of the enemy patterns in the shooting game.

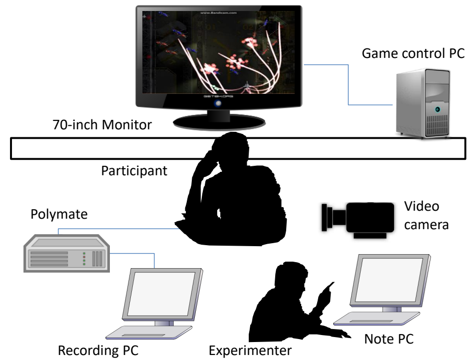


Figure 5. The experimental environment.

was over or cleared, the participants could confirm their score. They were instructed what behavior produced a high score, but it was difficult to carry out such behavior during the game. The participants acquired game playing skills through repeated practice. The game playing skill levels that we assumed are as follows.

- Basic: A player can clear the game stage.
- Senior: A player can get high scores in easy parts of the game stage.
- Expert: A player can find the way to get high scores in difficult parts, but he/she sometimes fail.
- Master: A player can get high scores in all parts of the game stage.

The easy parts are the first, fourth and seventh parts because a player can easily find the way to get high score. When a player can get high score in these parts, his/her skill is a senior level. The difficult parts are the second, third, fifth and sixth parts. The types of difficulties are different between the second and third parts and the fifth and sixth parts. In the second and third parts, a player has to deal with enemies of different behaviors in parallel. In the fifth and sixth parts, a player has to deal with enemies according to a difficult specific procedure. When a player can get high score in either type of parts, his/her skill is an expert level. When a player can get high score in all parts, his/her skill is a master level.

B. *Experimental set-up*

The experimental set-up is shown in Figure 5. Each participant sat in front of a 70-inch monitor that displayed the game. The participants used a joystick with two buttons for controlling the player character. A video camera was placed behind the participant to record his/her behavior and the game playing screen. The participant's voice was recorded using microphones of the video camera. Polymate was used to measure SCR and the electrocardiogram. SCR was measured with electrodes attached to the first and third fingers of the participant's non-dominant hand. The electrocardiogram was measured by connecting electrodes with paste to the participant's left side, the center of the chest, and both ears for ground and reference. The experimenter sat out of view of the participant and made notes about the participant's behavior and his/her subjective feeling of difficulty during each part. The experimenter was an expert at the game with experience teaching novice players how to get a high score and thus knew the points where to get a higher score and where is difficult for novice players.

C. *Participants*

The participants in the experiment were 21 male undergraduate students between the ages of 18 and 25 (with a mean age of 21.7). We eliminated the data of four participants, two of whom did not acquire the skills because their game scores did not increase throughout the experiment, and for two of whom we failed to measure the physiological indices because the electrodes detached. Therefore, the data of 17 participants were used for analysis. The participants repeatedly played the game in the experiment a total of 50 times; the total duration of the experiment was about 240 minutes. Therefore, we could obtain 850 game playing samples in all for analyses.

D. *Procedure of the experiment*

The experiment was divided into two sessions in the middle, and the participants took a long rest (15 minutes or more) between sessions because playing the game and acquiring the skills required heavy concentration. Each participant was briefly instructed on the experimental procedure. Electrodes for measuring SCR and the LF/HF electrocardiogram values were then attached to the participant's left hand and chest. The participant then played practice games twice. After a two-minute relaxation period, the experimenter started the video cameras and began recording the physiological indices, whereupon the participant began an experimental session. The participant played the game until it was over or cleared, then relaxed for 30 seconds. Moreover, the participants rested for 3 minutes every 10 games, during which time the experimenter scored the participant's feeling of difficulty subjectively and interviewed him/her about this sentiment at each part of the stage. In the first session of the experiment, the participants played 30 games because many games ended in the middle of the stage in the first session.

E. *Analysis*

We analyzed the data obtained from the experiment. We mainly focused on the data of physiological indices.

1) The relationship between participant's statuses and the physiological indices: The data obtained in the experiment are explained below.

1) The game score of each game play
2) The game score of each part of the stage
3) The feeling of difficulty as scored by the experimenter

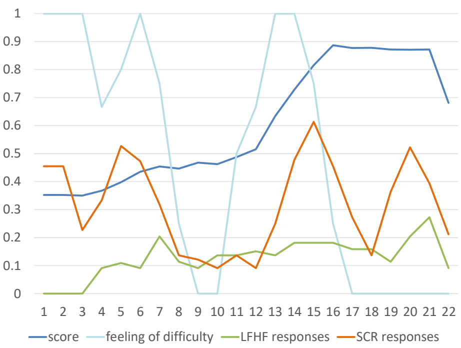


Figure 6. An example of analyzed data.

4) The physiological indices during the game play

We used 2 as the task performance and 3 as the feeling of difficulty in our model. In this analysis, we used the simple moving average (calculated from data for the previous five games and shifted by two games) for the data from 2 and 4 . The data from 3 served as categorical data to identify the feeling of difficulty. For example, the first data point was calculated from the data for games 1–5, the second data point from the data for games 3–7, and so on. Figure 6 shows an example of the analyzed data. A horizontal axis shows the data number of the simple moving average, and a vertical axis shows the ratio of the each value in each part. The task performance was assigned the value of the score for each part, divided by the maximum score for that part. The feeling of difficulty took on the values +1, i.e., felt difficulty at that part of the stage, and 0, i.e., did not feel difficulty at that part of the stage. The physiological responses were measured as the total time over the threshold (LF/HF: 3.0, SCR: 15.5) divided by the total time of that part.

We did not analyze the eighth part because we could not objectively segment the battle with the boss enemy. Thus, for this section we analyzed the first seven parts.

2) The relationship between the feeling of difficulty and the physiological indices: We calculated Spearman's rank correlation coefficients between the values of the task performance (TP) and the values of physiological responses (LF/HF and SCR), and those between the values of the feeling of difficulty (FD) and the values of physiological responses. The results are shown in Table I. In the second part, there is a weak negative correlation between the values of the task performance and the SCR responses, meaning that the scores themselves were not so strongly related to the participants' mental states. On the other hand, there are weak positive correlations between the values of the feeling of difficulty and the SCR responses in three out of the seven parts, and moderate correlations in three other parts. This confirms that there is a relationship between the feeling of difficulty and SCR.

We also performed Mann-Whitney U tests to compare the physiological responses when participants felt difficulty with the task and those when they did not. The numbers of samples are as follows (part number - number of data points where participants felt difficulty : number of other data points); the first part - 250 : 597, the second part - 370 : 472, the third part - 127 : 704, the fourth part - 575 : 275, the fifth part - 449 : 274, the sixth part - 248 : 417, and the seventh part - 216 : 395. The results are shown in Table II. In this, only the p-values under 0.1 are shown easy to see. The hyphen ("-") means that the p-value is over 0.1. There is a significant difference in the LF/HF responses only in the second part. On the other hand, there are significant differences in SCR responses in all the parts. These results indicate that SCR is related to the feeling of difficulty.

TABLE I. CORRELATION COEFFICIENTS BETWEEN THE VALUES OF THE TP AND THE PHYSIOLOGICAL RESPONSES, AND THOSE BETWEEN THE VALUES OF THE FD AND THE PHYSIOLOGICAL RESPONSES.

	LF/HF and TP	SCR and TP	LF/HF and FD	SCR and FD
1st part	-0.064	-0.189	0.042	0.354
2nd part	-0.084	-0.212	0.065	0.462
3rd part	-0.063	-0.029	-0.035	0.136
4th part	-0.018	-0.041	-0.011	0.419
5th part	0.04	0.07	0.017	0.318
6th part	-0.006	-0.065	0.036	0.5
7th part	-0.045	-0.009	-0.042	0.317

TABLE II. RESULTS OF MANN-WHITNEY U TESTS BETWEEN DATA IN THE NON-FD AND THE FD.

	LF/HF between FD and non-FD	SCR between FD and non-FD
1st part	-	0.0000
2nd part	0.0594	0.0000
3rd part	-	0.0001
4th part	-	0.0000
5th part	-	0.0000
6th part	-	0.0000
7th part	-	0.0000

In our previous study, we reported that SCR relatively reflected intrinsic stress (concentrating on the task, considering it, and so on) and that LF/HF relatively reflected the effects by extrinsic stimuli [21]. We expected that SCR would respond when the participant considered how to destroy enemies and what a bad point was in prior game plays, and therefore that we would find a relationship between the feeling of difficulty and SCR. However, we expected that the LF/HF would respond when the participant encountered impressive game events such as getting a high score or making a mistake. Impressive game events occurred independently of task performance and the feeling of difficulty. Therefore, we could not find a relationship with LF/HF.

3) The differences of the physiological indices between sections where task performance rapidly increased and others: If SCR reflects intrinsic stress, there is a link between SCR responses and changes of task performance, especially in the parts of the task where it is important to plan how to destroy enemies and get high scores. On the other hand, if LF/HF reflects the effects of external stimuli, LF/HF responses change whether the participants' technique is optimized or not. We thus expect that the values of the LF/HF responses and the SCR responses in sections where task performance rapidly increased are important for assessing the skill acquisition process in some parts of the stage. We refer to a section where task performance rapidly rose as a "Rapidly Increasing Section" (RIS), which is defined as a section in which task performance increases by more than 0.1 for a continuous two or more sections.

We compared the LF/HF responses and the SCR responses between the values in RISs and the other sections. We expected

TABLE III. RESULTS OF THE MANN-WHITNEY U TEST BETWEEN DATA OF LF/HF IN NON-RIS AND THAT IN RIS.

	The whole part	The 1st half	The 2nd half
1st part	-	-	0.0575
2nd part	0.0020	0.0021	0.0474
3rd part	-	-	-
4th part	-	-	-
5th part	0.0835	-	-
6th part	-	-	-
7th part	-	-	0.0491

TABLE IV. RESULTS OF THE MANN-WHITNEY U TEST BETWEEN DATA OF SCR IN NON-RIS AND THAT IN RIS.

	The whole part	The 1st half	The 2nd half
1st part	0.0004	0.0031	0.0004
2nd part	-	-	-
3rd part	0.0094	0.0012	-
4th part	-	-	-
5th part	0.0135	-	0.0000
6th part	-	-	-
7th part	0.0011	0.0147	0.0009

TABLE V. RESULTS OF THE MANN-WHITNEY U TEST BETWEEN DATA OF LF/HF IN NON-PS AND THAT IN PS.

	The whole part	The 1st half	The 2nd half
1st part	-	0.0066	0.0905
2nd part	-	-	-
3rd part	-	-	-
4th part	0.0246	0.0576	0.0957
5th part	-	-	-
6th part	0.0498	0.0829	0.0983
7th part	0.0821	-	0.0338

TABLE VI. RESULTS OF THE MANN-WHITNEY U TEST BETWEEN DATA OF SCR IN NON-PS AND THAT IN PS.

	The whole part	The 1st half	The 2nd half
1st part	0.0900	-	0.0822
2nd part	-	-	-
3rd part	0.0082	0.0636	0.0231
4th part	-	-	-
5th part	-	-	-
6th part	-	0.0467	0.0898
7th part	-	-	-

that the meaning of the physiological responses was different between the first and second halves of the part of the stage. The physiological responses in the first half show whether the participant had a concrete plan for scoring in advance. The physiological responses in the second half show whether the participant was used to performing the task in the part. We thus calculated the values of the LF/HF responses and the SCR responses in the whole game part and in both halves. We performed Mann-Whitney U tests for each set of values of the physiological indices. The numbers of samples are as follows (part number - number of data points in the rapid increasing section : number of other data points); the first part - 139 : 275, the second part - 128 : 263, the third part - 145 : 246, the fourth part - 94 : 297, the fifth part - 93 : 292, the sixth part - 134 : 250, and the seventh part - 101 : 271. The results are shown in Table III and Table IV. In the tables, only the p-values under 0.1 are shown easy to see. The hyphen ("-") means that the p-value is over 0.1. In the LF/HF responses for the whole game part, there is a significant difference in the second part and a marginally significant difference in the fifth part. In the SCR responses for the whole game part, there are significant differences in the first, third, fifth, and seventh parts.

Meanwhile, for the physiological responses in the first and second halves of the game part, the patterns of significant differences are different. For example, in the fifth part, there is no significant difference in the first half of the SCR responses. The task of the fifth part is to destroy a mid-boss enemy, and since the behavior of the mid boss becomes complex over time, it is reasonable that the player executes trial and error in the second half. In the case of the second part, there is a significant difference only in the LF/HF responses. In the second part, the patterns of the enemies change in the middle. Since the player has to respond the changes reactively, it is reasonable that the player takes care about extrinsic stimuli.

To sum up, the results suggest that we may distinguish the task features in the skill acquisition process using the transitions of task performance and the different patterns of the physiological responses.

4) The differences of the physiological indices between sections where task performance decreased and others: The RIS is a good example in the skill acquisition process. However, task performance sometimes decreases because the learner tries to optimize his/her behavior to the task. During optimization, the learner mainly pays attention to how to respond to the external stimuli. We thus expect that the values of the LF/HF responses and the SCR responses in such a section will fall. We refer to the section where the task performance falls into a "Plateau Section" (PS), defined as a section in which task performance stays constant or decreases continuously for two or more sections.

We compared the values of the LF/HF responses and the SCR responses between the plateau sections and the other sections. In the same way as with the RISs, we calculated the values of the LF/HF responses and the SCR responses in the whole game part and the two halves of the game part. We performed Mann-Whitney U tests on each set of values of physiological indices. The numbers of samples are as follows (part number - number of data points in the plateau section : number of other data points); the first part - 242 : 172, the second part - 218 : 173, the third part - 228 : 163, the fourth part - 256 : 135, the fifth part - 262 : 129, the sixth part - 243 : 141, and the seventh part - 249 : 123. The results are shown in the Table V and Table VI. In the tables, only the p-values under 0.1 are shown easy to see. The hyphen ("-") means that the p-value is over 0.1. Unlike the previous analysis, there is a major difference in the patterns of the significant differences in the whole and the two halves of the plateau section. In the fourth and sixth parts, there are significant differences for the whole plateau sections. On the other hand, in the first and seventh parts, there are significant differences in the first or second half of the plateau sections. In addition, we can find more differences in the LF/HF responses than in the SCR responses, suggesting that the learners may try to optimize their behavior in the short subtasks in the first and seventh parts in the plateau section. This is supporting evidence that a situation may arise where task performance does not increase even when the learner acquires sub-skills.

In this study, we know the maximum game scores in each part of the stage. Therefore, we can determine that learners try to optimize their behavior or that they are performing tasks through inertia. The plateau section is important because it is the preparation phase for the skill acquisition process. However, in terms of general skill tasks, we cannot know the maximum performance of the endeavor. Hence, we propose a method to identify plateau sections using task performance and the learner's mental state.

F. The estimation of participant's statuses using the physiological indices

As mentioned above, we can find some significant differences in the sections with the feeling of difficulty, the rapid increasing sections, and the plateau sections, which suggests that the physiological indices are clues to estimate the participant's status in skill acquisition. Of course, other clues are needed for accurate estimation, such as task structures, the basic abilities of the learners, and the details of the mental states of the learners. However, as a first step, we try to estimate the three states in the skill acquisition: whether the learner feels difficulty in a particular part of the task, whether a particular part of the task is a rapid increasing section or not, and whether a particular part of the task is a plateau section or not. To classify the states, we used Support Vector Machine (SVM) Modeling with radial basis function kernels. SVMs are supervised learning models. SVMs have been introduced for solving pattern recognition problems [22][23]. The SVM learning algorithm builds a model that assigns new cases to a certain category or another category. SVM is a non-probabilistic binary linear classifier. SVM uses a kernel function which maps the given data into a different space; the separations can be made even with very complex boundaries. The hyperplane algorithm is a way to create nonlinear classifier by applying the kernel trick to maximum margin hyperplanes.

We calculated the values of the LF/HF responses and the SCR responses in the whole game part and the two halves. In addition, we also calculated the values of the mean and the standard deviation of LF/HF and SCR. To estimate the feeling of difficulty, the values were calculated for each trial of the experiment. When estimating the rapid increasing section and the plateau section, the values of the simple moving average (calculated from data for the previous five games and shifted by two games) were calculated. The 18 variables were used for SVM with radial basis function kernels. The variables were selected by the stepwise method. In this analysis, we performed leave-one-game-part-data-out cross validation for SVM classification using data from which the data of one game part were removed. The SVM model was applied to the data for each game part and the F-measure and accuracy were calculated. In the following tables, each row shows the F-measure and accuracy resulting from SVM classification using data from which the data for the game part for the row were removed. For example, the data in the second row and the fourth column of the table shows the result of applying SVM model to the data of the fourth part, calculated using the data from which the data for the second part were removed. Therefore, the diagonal values are the results of the cross validations. In the following table, the diagonal values are displayed in a bold typeface.

TABLE VII. F-MEASURES OF SVM CLASSIFICATIONS FOR THE ESTIMATION OF THE FEELING OF DIFFICULTY.

	1st	2nd	3rd	4th	5th	6th	7th
1st part	**0.523**	0.708	0.365	0.802	0.797	0.720	0.627
2nd part	0.610	**0.634**	0.358	0.810	0.802	0.717	0.638
3rd part	0.616	0.713	**0.293**	0.817	0.811	0.722	0.626
4th part	0.628	0.729	0.352	**0.741**	0.805	0.749	0.647
5th part	0.661	0.711	0.366	0.826	**0.723**	0.742	0.646
6th part	0.611	0.684	0.368	0.804	0.807	**0.656**	0.621
7th part	0.599	0.706	0.365	0.826	0.813	0.741	**0.569**

TABLE VIII. ACCURACIES OF SVM CLASSIFICATIONS FOR THE ESTIMATION OF THE FEELING OF DIFFICULTY.

	1st	2nd	3rd	4th	5th	6th	7th
1st part	**0.694**	0.757	0.740	0.749	0.736	0.789	0.691
2nd part	0.758	**0.702**	0.750	0.756	0.748	0.788	0.709
3rd part	0.750	0.752	**0.657**	0.760	0.754	0.780	0.679
4th part	0.747	0.765	0.708	**0.675**	0.734	0.800	0.700
5th part	0.780	0.751	0.724	0.770	**0.651**	0.791	0.697
6th part	0.759	0.736	0.752	0.749	0.754	**0.743**	0.694
7th part	0.744	0.752	0.732	0.772	0.757	0.798	**0.604**

1) The estimation of the feeling of difficulty: Table VII and Table VIII show the results of SVM classifications. The numbers of samples are same as in the section of "The relationship between the feeling of difficulty and the physiological indices". The diagonal values show results when the data not used to calculate the SVM model. As a result, the values of F-measure were over 0.6 in most cases and over 0.7 in 25/49 cases. The values of the accuracy were over 0.7 when the data used to calculate the SVM model were classified and over 0.65 when the data not used to calculate the SVM model (except the data in the seventh part) were classified. The results suggest the possibility of estimating the feeling of difficulty using physiological indices.

From the results, we can find that the F-measures are relatively low in the data of the third part, but the accuracies are not so low. This means that it is hard to classify the data points in the rapid increasing section. The enemy pattern in the third part was a little different from other part; two enemies that has high score parallelly appeared with another type of enemies in the part. Since, in this part, there are some strategies to get score, the stress of the participant was changed depending on the adopted strategy. In addition, depending on the adopted strategy, the feeling of difficulty of the participant was changed, but it was difficult to identify the adopted strategy. Therefore, we expect that the data of the feeling of difficulty as scored by the experimenter included some errors.

2) The estimation of the RIS: Table IX and Table X show the results of SVM classifications. The numbers of samples are same as in the section of "The differences of the physiological indices between sections where task performance rapidly increased and others". The diagonal values show results when the data not used to calculate the SVM model. As a result, the values of the F-measure and accuracy were low when the data not used to calculate the SVM model were classified. The values of the F-measure and the accuracy were over 0.5 and over 0.65, respectively, when the data used to calculate the SVM model were classified. The results suggest that the generalization capability was low, but that there is a

TABLE IX. F-MEASURES OF SVM CLASSIFICATIONS FOR THE ESTIMATION OF THE RIS.

	1st	2nd	3rd	4th	5th	6th	7th
1st part	**0.417**	0.573	0.553	0.535	0.638	0.634	0.655
2nd part	0.606	**0.401**	0.616	0.523	0.600	0.618	0.625
3rd part	0.584	0.556	**0.406**	0.535	0.651	0.634	0.639
4th part	0.596	0.587	0.592	**0.283**	0.642	0.618	0.675
5th part	0.615	0.578	0.614	0.568	**0.314**	0.644	0.658
6th part	0.601	0.586	0.601	0.531	0.624	**0.305**	0.702
7th part	0.589	0.599	0.563	0.576	0.591	0.658	**0.331**

TABLE X. ACCURACIES OF SVM CLASSIFICATIONS FOR THE ESTIMATION OF THE RIS.

	1st	2nd	3rd	4th	5th	6th	7th
1st part	**0.534**	0.665	0.632	0.693	0.760	0.708	0.782
2nd part	0.674	**0.435**	0.627	0.655	0.724	0.672	0.728
3rd part	0.684	0.652	**0.565**	0.693	0.770	0.711	0.766
4th part	0.696	0.680	0.683	**0.430**	0.754	0.716	0.790
5th part	0.686	0.634	0.662	0.731	**0.476**	0.729	0.780
6th part	0.676	0.650	0.650	0.670	0.731	**0.503**	0.806
7th part	0.667	0.655	0.642	0.706	0.734	0.729	**0.522**

TABLE XI. F-MEASURES OF SVM CLASSIFICATIONS FOR THE ESTIMATION OF THE PS.

	1st	2nd	3rd	4th	5th	6th	7th
1st part	**0.462**	0.584	0.659	0.691	0.738	0.696	0.803
2nd part	0.684	**0.538**	0.656	0.676	0.733	0.720	0.746
3rd part	0.647	0.541	**0.495**	0.712	0.759	0.682	0.736
4th part	0.675	0.637	0.701	**0.563**	0.727	0.754	0.761
5th part	0.683	0.656	0.679	0.745	**0.650**	0.723	0.809
6th part	0.664	0.589	0.606	0.664	0.734	**0.636**	0.777
7th part	0.695	0.602	0.618	0.748	0.765	0.747	**0.620**

TABLE XII. ACCURACIES OF SVM CLASSIFICATIONS FOR THE ESTIMATION OF THE PS.

	1st	2nd	3rd	4th	5th	6th	7th
1st part	**0.488**	0.606	0.642	0.657	0.675	0.641	0.758
2nd part	0.640	**0.504**	0.619	0.634	0.662	0.661	0.688
3rd part	0.626	0.558	**0.504**	0.668	0.675	0.646	0.680
4th part	0.616	0.601	0.645	**0.504**	0.668	0.688	0.702
5th part	0.633	0.609	0.616	0.688	**0.545**	0.656	0.753
6th part	0.597	0.542	0.560	0.611	0.650	**0.549**	0.715
7th part	0.640	0.593	0.575	0.680	0.675	0.682	**0.530**

possibility of estimating the rapid increasing section using the data used to calculate the SVM model. Of course, the values of accuracy when the data used to calculate the SVM model were classified were not sufficient in some parts. We have to consider more details in the participant's mental states and the conditions of the task.

3) The estimation of the PS: Table XI and Table XII show the results of SVM classifications. The numbers of samples are same as in the section of "The differences of the physiological indices between sections where task performance decreased and others". The diagonal values show results when the data not used to calculate the SVM model. As a result, the values of the F-measure and accuracy were not so high when the data not used to calculate the SVM model were classified. The values of F-measure and accuracy in the plateau section were lower than those in the rapid increasing section. One reason is that situations in which learners try to optimize their behavior and situations in which they are performing tasks through inertia are mixed. More detailed teacher data are needed to improve the classification.

V. DISCUSSION

This study aimed to develop a method to estimate the learner's skill level based on task performance and his/her mental states. To achieve this, we conducted an experiment to obtain data from the subjective reports of a feeling of difficulty and the physiological indices during the skill acquisition process. We then confirmed the relationship between the subjective reports of the feeling of difficulty and the physiological indices. In addition, we suggested an approach to identify the statuses of the skill acquisition process, e.g., the learner feels difficulty and the task performance is rapidly increasing, via the measured physiological indices. Concretely, we classified the trials with a feeling of difficulty in the experiment using SVM. As a result, the values of accuracy were over 0.7 when the data used to calculate the SVM model were classified and over 0.65 when the rest of the data not used to calculate the SVM model were classified. When we classified the trials in the rapid increasing section, the values of accuracy were over 0.65 when the data used to calculate the SVM model were classified. However, when we classified the trials in the plateau section, the values of accuracy were low. Although we have to refer to the teacher data for more accurate classification, we suggest that there is a possibility of estimating the statuses of the skill acquisition process.

In the results of the Mann-Whitney U tests, there is a different tendency between the results for the feeling of difficulty and the results of the rapid increasing and plateau sections. In the results of the feeling of difficulty, the SCR response is significantly different in all parts of the game but the LF/HF response is not. On the other hand, in the results of the rapid increasing section and the plateau section, the SCR response is not significantly different in certain parts of the game, while the LF/HF response is significantly different in such parts. During the skill acquisition process, learners execute actions by trial and error, receive feedback, and improve their behavior based on subjective and objective evaluations. The learners feel difficulty when they evaluate their behavior after finishing the task or sub-task and thus are not influenced by the extrinsic stimuli in the task, i.e., their LF/HF values are not changed. When the learners improve their behavior in real time, they have to pay attention to the extrinsic stimuli in the task depending on the statuses of the task, e.g., the patterns of the enemies and whether they already have a plan to score. The patterns of the significant differences in the SCR and LF/HF responses reflect the statuses of the task.

A comparison of the results of the Mann-Whitney U tests and those of the SVM classification shows that the physiological responses showing significant differences in the Mann-Whitney U tests do not directly contribute to the SVM classification. One reason is that the distributions of the physiological responses overlap between the statuses. For example, the results for the third part in the feeling of difficulty show no difference from the others, but it is difficult to classify the data for the third part using the SVM. Conversely, it is not so difficult to classify the data for the sixth part in the rapid increasing section using the SVM, but there is no significant difference in the physiological responses. These suggest that the skill acquisition of the skill task has complex processes related to different factors and only one particular parameter does not suffice to show the status of the skill acquisition. Meanwhile, we suggest that the multiple physiological indices are important clues to the status of the skill acquisition process.

The most important contribution of this study is that we experimentally analyzed the effects of the learner's mental states based on the physiological indices. In a traditional skill acquisition model, the impacts of the learner's mental states are conceptually proposed but not confirmed objectively. Of course, our study has some limitations. The most serious limitation was that we could not achieve sufficient F-measure and accuracy in the SVM classification when the data not used to calculate the SVM model were classified. One reason is that the important points for acquiring the skill are different for each part of the game. For example, in the third part, participants need to pay attention to the different types of enemies in parallel. For accurate classification, it is an important piece of information, which enemy was paid attention to, but such a clue could not be included in the SVM classification. Another limitation was the segmentation of the task. In this research, we segmented the game into eight parts (including the final boss part). However, the segmentations were not sufficient for accurate classification, which is a task awaiting future work on realizing a skill acquisition support system.

VI. Conclusion and Future Work

The purpose of this study was to experimentally investigate whether the statuses of the skill acquisition process could be estimated by the physiological indices used to estimate the learner's mental states. For this purpose, we conducted an experiment to obtain the data of the physiological indices and a subjective report of the feeling of difficulty during the skill acquisition task. As a result of the analysis, we confirmed the relationship between the participant's statuses (feeling of difficulty, the rapid increasing section, and the plateau section) and the physiological indices. In addition, we classified the trials having a feeling of difficulty in the experiment using SVM. Overall, we could show that the physiological indices are helpful clues to estimate the status of the skill acquisition process. In future research, we will obtain more teacher data for machine learning techniques and analyze the different features of the skill acquisition task. In addition, we will extend the framework of experiments to different tasks that include the different features of the skill acquisition task.

VII. Acknowledgement

This research is supported by Grant-in-Aid for Young Scientists (B) (KAKENHI No. 16K21113), and Grant-in-Aid for Scientific Research on Innovative Areas (KAKENHI No. 26118002) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

[1] Y. Ohmoto, T. Matsuda, and T. Nishida, "Experimentally analyzing the skill acquisition model using task performance and physiological indices," in COGNITIVE 2017: The ninth International Conference on Advanced Cognitive Technologies and Applications. IARIA, 2017, pp. 44–49.

[2] S. E. Dreyfus and H. L. Dreyfus, "A five-stage model of the mental activities involved in directed skill acquisition," DTIC Document, Tech. Rep., 1980.

[3] J. Tranquillo and M. Stecker, "Using intrinsic and extrinsic motivation in continuing professional education," Surgical neurology international, vol. 7, no. Suppl 7, 2016, p. S197.

[4] K. Kraiger, J. K. Ford, and E. Salas, "Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation," Journal of applied psychology, vol. 78, no. 2, 1993, p. 311.

[5] T. R. Mitchell, H. Hopper, D. Daniels, J. George-Falvy, and L. R. James, "Predicting self-efficacy and performance during skill acquisition," Journal of Applied Psychology, vol. 79, no. 4, 1994, p. 506.

[6] F. Gobet and P. Chassy, "Expertise and intuition: A tale of three theories," Minds and Machines, vol. 19, no. 2, 2009, pp. 151–180.

[7] E. A. Day, W. Arthur Jr, and D. Gettman, "Knowledge structures and the acquisition of a complex skill," Journal of applied psychology, vol. 86, no. 5, 2001, p. 1022.

[8] J. Langan-Fox, K. Armstrong, N. Balvin, and J. Anglim, "Process in skill acquisition: Motivation, interruptions, memory, affective states, and metacognition," Australian Psychologist, vol. 37, no. 2, 2002, pp. 104–117.

[9] P. M. Fitts and M. I. Posner, "Human performance," 1967.

[10] J. R. Anderson, "Acquisition of cognitive skill," Psychological review, vol. 89, no. 4, 1982, p. 369.

[11] R. M. Shiffrin and W. Schneider, "Controlled and automatic human information processing: Ii. perceptual learning, automatic attending and a general theory," Psychological review, vol. 84, no. 2, 1977, p. 127.

[12] C. Konrad, G. Schupfer, M. Wietlisbach, and H. Gerber, "Learning manual skills in anesthesiology: is there a recommended number of cases for anesthetic procedures?" Anesthesia & Analgesia, vol. 86, no. 3, 1998, pp. 635–639.

[13] F. E. Ritter and L. J. Schooler, "The learning curve," International encyclopedia of the social and behavioral sciences, vol. 13, 2001, pp. 8602–8605.

[14] T. P. Grantcharov and P. Funch-Jensen, "Can everyone achieve proficiency with the laparoscopic technique? learning curve patterns in technical skills acquisition," The American Journal of Surgery, vol. 197, no. 4, 2009, pp. 447–449.

[15] J. T. Patterson, A. Hart, S. Hansen, M. J. Carter, and D. Ditor, "Measuring investment in learning: Can electrocardiogram provide an indication of cognitive effort during learning?" Perceptual and motor skills, vol. 122, no. 2, 2016, pp. 375–394.

[16] A. D. Walker, E. R. Muth, F. S. Switzer III, and P. J. Rosopa, "Predicting team performance in a dynamic environment: A team psychophysiological approach to measuring cognitive readiness," Journal of Cognitive Engineering and Decision Making, vol. 7, no. 1, 2013, pp. 69–82.

[17] P. Parsinejad, "Inferring mental workload changes of subjects unfamiliar with a touch screen game through physiological and behavioral measurements," Ph.D. dissertation, Northeastern University, 2016.

[18] E. F. Bartholomew, F. Martini, and W. B. Ober, Essentials of anatomy & physiology. Benjamin Cummings, 2007.

[19] K. Hugdahl, Psychophysiology: The mind-body perspective. Harvard University Press, 1995.

[20] B. C. Lacey and J. I. Lacey, "Two-way communication between the heart and the brain: Significance of time within the cardiac cycle," American Psychologist, vol. 33, no. 2, 1978, p. 99.

[21] Y. Ohmoto, S. Takeda, and T. Nishida, "Distinction of intrinsic and extrinsic stress in an exercise game by combining multiple physiological indices," in Proceedings of the seventh International Conference on Games and Virtual Worlds for Serious Applications (VS-Games). IEEE, 2015, pp. 1–4.

[22] V. Vapnik, "The support vector method of function estimation," Nonlinear modeling: Advanced black-box techniques, vol. 55, 1998, p. 86.

[23] N. Cristianini and J. Shawe-Taylor, An introduction to support vector machines and other kernel-based learning methods. Cambridge University Press, 2000.