Exploring Patterns of Skill Gain and Loss on Long-term Training and Non-training in Rhythm Game

Ayane Sasaki[1], Mio Matsuura[1], Masaki Matsubara[2], Yoshinari Takegawa[1], and Keiji Hirata[1]

[1] Future University Hakodate
[2] University of Tsukuba

Abstract. The objective of this study is to categorize patterns of skill loss following skill gain, in order to develop a predictive model for skill retention in music games. The experiment was conducted using songs from the web-based music game “Sparebeat.” Participants were instructed to train daily on a piece of music slightly more challenging than their current skill level until they achieved a specified level of proficiency. Following this, participants took a break from training for at least one week, and their scores were recorded when they played the music immediately after the non-training phase. By analyzing the changes in scores during both the skill gain and loss phases, we identified three distinct patterns of skill loss.

Keywords: Educational Technology, Human Computer Interaction

1 Introduction

In recent years, advances in HCI technology have enabled the proposal of numerous learning support systems to assist learners in acquiring skills involving physical movements, such as tennis [1], golf [2–4], calligraphy [5, 6], playing musical instruments [7], and singing [8].

A common learning support framework involves adjusting the difficulty of skill gain based on the learner’s level, providing a sense of accomplishment during training and fostering motivation. For instance, bicycles equipped with training wheels enable inexperienced riders to train and eventually ride without assistance. The main challenges lie in determining how to modify the target skill’s difficulty and provide learners with environments that promote continued motivation.

While research has explored the cognitive aspects of skill gain, it is important to consider “skill loss,” the decline in acquired skills that occurs once a person stops training. Factors like individual differences in skill loss suggest that cognitive aspects of learners are involved in this process. By examining both skill gain and loss, we can develop a more accurate cognitive model of skill gain.

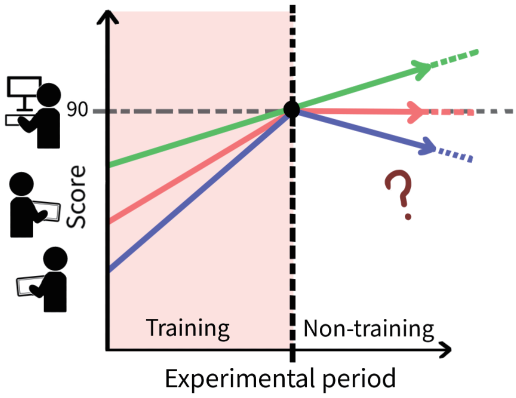

Fig. 1. Overview: our study conducts long-term experiment to explore skill gain and loss patterns through training and non-training periods

Previous studies have investigated memorization and forgetting in tasks that do not involve physical actions, such as memorizing words [9–13]. In sports science, on the assumption that a sportsperson trains daily, research has focused on the relationship between sleep and memory consolidation through motor skill learning [14–16].

Our study aims to clarify the relationship between skill gain and loss in music games for beginners, by observing score changes and timing judgment classifications during both the skill gain and loss phases (Fig. 1). We also seek to discover and classify patterns of skill loss.

A unique feature of our study is its long-term experimental design, as the gain and loss of skills involving physical movements require a certain length of time. For example, learning to ride a bicycle typically takes several days to weeks. In sports like tennis and golf, there is virtually no upper limit to the time required for skill gain. Acquired skills are not forgotten until a certain length of time has passed[3]. In our experiment, participants trained until they reached a specific music game score, with some requiring up to 50 days of training. After the training phase, subjects entered a skill loss phase, with some continuing the experiment for nearly 90 days. Observing gain and loss over such a long period is expected to yield essential data and findings.

Music games are excellent targets for experiments involving skill gain and physical movement. As games, they inherently motivate players to continue practicing, and players can engage in music games for extended periods without boredom. Music games require a certain level of skill and training to achieve a high score and offer a mechanism to consistently and stably assess a player's level of skill acquisition.

Towards identifying the relationship between skill gain and loss in music games, we conduct an exploratory study to find an appropriate hypothesis as the first step.

[3] Once a person is able to ride a bicycle, he or she will not completely forget the skill, although the skill level may deteriorate.

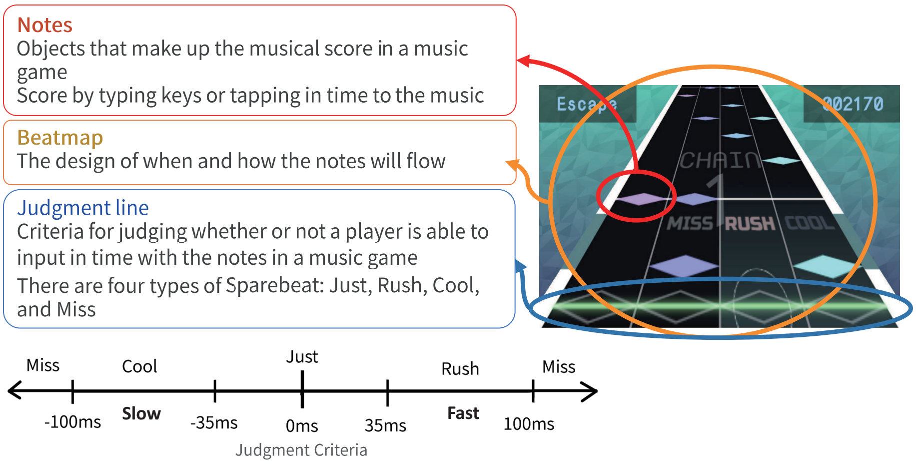

Fig. 2. Screenshot of music game 'Sparebeat' and criteria of judgement

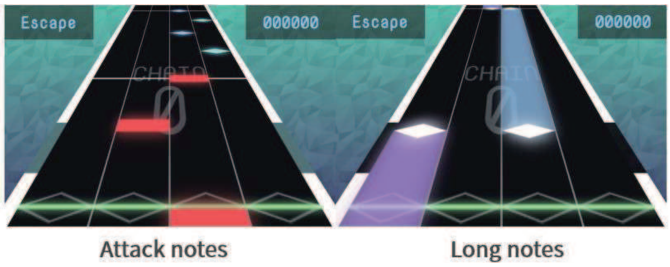

Fig. 3. Attack notes and Long notes

2 Materials and Methods

2.1 Music game Sparebeat

We used Sparebeat (Fig. 2), a music game simulator, for our experiments. It runs on web browsers and is playable on PCs, smartphones, and tablets. The playing screen consists of four black lanes with diamond-shaped notes in different colors moving towards a green line. Players must press the corresponding key as the notes cross the green line to earn points. Sparebeat has three types of notes with varying difficulty levels and display formats.

Music pieces for the experiment were selected based on each subject's skill level from a preliminary assessment. We chose pieces with a score of 650,000 to 700,000 points to ensure they were neither too easy nor too difficult, allowing us to measure skill gain effectively.

Sparebeat has four types of timing judgments for key presses: *Just, Rush, Cool, and Miss*. Each judgment depends on the accuracy of the player's timing when pressing the keys. As indicated by the criteria arrows at the bottom of Fig. 2, Just is correct, Rush is

Table 1. Judgment and score criteria

	Just	Rush	Cool	Miss
Normal notes	100%	50%	50%	0%
Long notes	100%	50%	50%	0%
Attack notes	200%	100%	100%	0%

Table 2. Experimental period

	Training phase	Non-training phase
Subject 1	16 days	90 days
Subject 2	10 days	58 days
Subject 3	10 days	58 days
Subject 4	34 days	63 days
Subject 5	50 days	50 days

fast, Cool is slow, and Miss is anything that does not fall into any of these categories. In addition to the "Normal notes" shown in Fig. 2, there are also "Long notes" and "Attack notes" (Fig. 3). The scoring system of Attack notes is different from that of Normal notes, and as shown in Table 1, the scoring is twice that of Normal notes.

The maximum score for any piece in Sparebeat is 1,000,000 points. The score per note varies depending on the piece and is calculated by dividing the full score by the total number of notes in the piece.

2.2 Participants

Five subjects participated, ranging from beginner to intermediate university and graduate students who had played music games as a hobby. None of the subjects had played Sparebeat before. They played the game on personal devices throughout the experiment.

2.3 Instructions for subjects

Subjects were asked to play their assigned piece once, train for 10 to 20 minutes, and then play it again. They trained daily, following a training set format. Once subjects consistently scored over 900,000 points, they entered a non-training phase during which they did not play the game. After this phase, they played their assigned piece once more, and their scores were recorded. This non-training set was repeated as necessary.

The threshold for suspending practice was determined to be 900,000 points due to experience and the results of when the Just, Rush, Cool, and Miss percentages exceed 900,000 points, which are discussed later in Section 3.1.

3 Results

3.1 Training phase

The duration of the training phase and the duration of the non-training phase for each subject are shown in Table 2. The length of the training phase and non-training phase

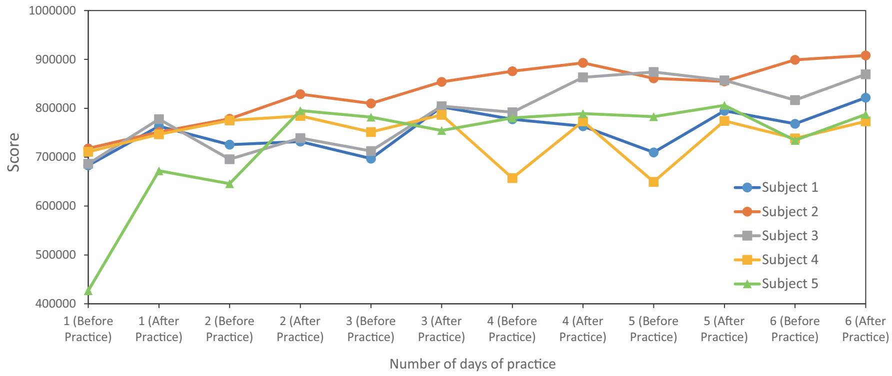

Fig. 4. Score transitions from Day 1 to Day 6 of training phase.

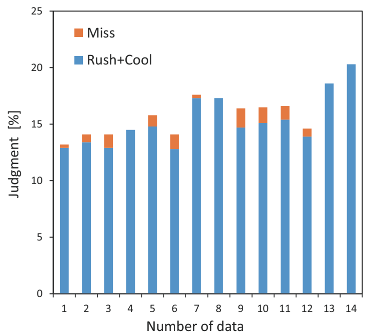

Fig. 5. Breakdown of Rush, Cool, and Miss ratio over 900,000 points among all subjects' performance

was different for each subject. The pre-training and post-training scores for each of the five subjects from the first day to the sixth day are shown in Fig. 4.

The graph in Fig. 4 shows that the post-training score is higher than the score on the first day of training on all training days. In addition, the pre-training score tends to be lower than the post-training score on a given training day. It can be said that the player generally improves with training. Although there were individual differences, the pre-training score was lower than the previous day's post-training score on the seventh day and beyond as well, but the score gradually increases with each training session.

During the training phase, the scores exceeding 900,000 are listed in descending order among the results that include all subjects before and after training, and the breakdown of Rush, Cool, and Miss in that data is shown in Fig. 5. Since the distribution of Rush and Cool scores is the same, the distribution of the percentage of the sum of Rush and Cool scores and the percentage of Miss scores is shown. From this figure, it

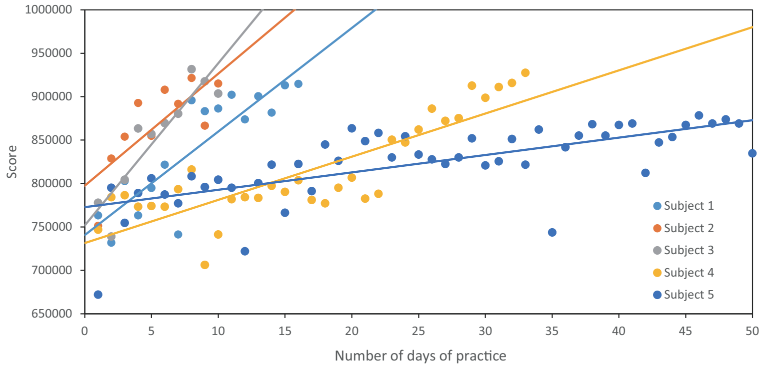

Fig. 6. Scatter plot of post-training scores in training phase

Table 3. Linear approximation of scores in training phase

	During training			Volatility	
	Slope	Intercept	Decision coefficient	Slope	Intercept
Subject 1	11,925	740,615	0.74	-4.9×10^{-4}	0.018
Subject 2	12,902	797,598	0.59	-7.6×10^{-3}	0.062
Subject 3	18,681	751,760	0.82	-3.0×10^{-3}	0.033
Subject 4	4,970	731,551	0.72	6.9×10^{-4}	-0.007
Subject 5	2,003	772,762	0.48	-5.2×10^{-4}	0.040

can be seen that the ratio of Rush and Cool must be approximately 15% or less to exceed 900,000 points.However, even when the ratio is larger than 15%, the score exceeds 900,000 points as long as the Miss ratio is approximately less than 1%.

During the training phase, we focus only on the post-training scores in order to investigate the evolution of scores until the skill is mastered. Fig. 6 shows a scatter plot of the post-training scores only. Table 3 shows the slope, coefficient of determination, etc. when a linear approximation is applied.

Subjects 1 to 3, who had trained for 10 to 16 days, had a slope of more than 10,000, indicating relatively rapid progress. Subjects 4 and 5 had a slope of less than 5,000. The training phases were 34 days, 50 days, and more than one month, respectively, meaning that progress was gradual, as it took time for these subjects to reach a certain skill level. The rate of change indicates how much score had changed when the score on a given day of the training phase was compared to that of the previous day. Then, the rate of change for each of the days up to the time when a subject stopped practicing is made into a regression line, and the slopes are shown in Table 3. From this, we can see that the slope is negative for all subjects except subject 4, indicating that the fluctuation of the score becomes smaller as training is repeated. It is possible that after a certain level of progress, the growth of the score nearly levelled off, and the score stabilized.

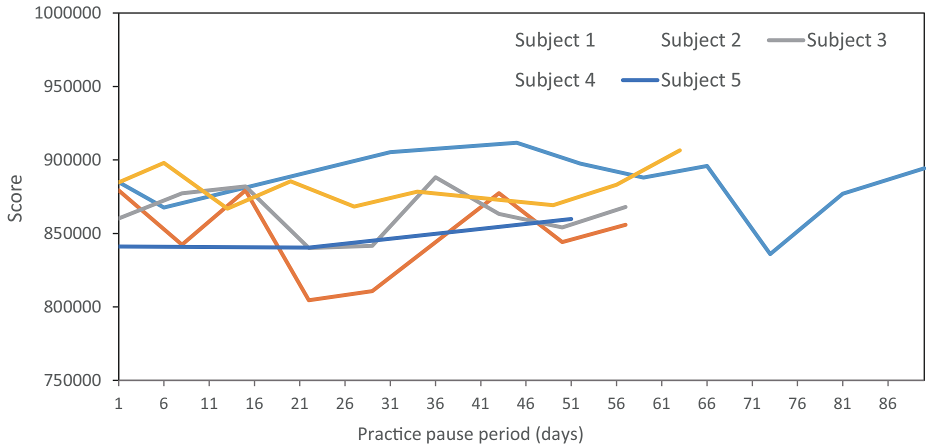

Fig. 7. Score transitions in non-training phase

Table 4. Linear approximation of scores in non-training phase

	Times	Slope	Intercept	Decision coefficient
Subject 1	10 times	-1,529	894,217	0.045
Subject 2	8 times	-1,834	857,464	0.023
Subject 3	9 times	-466	866,201	0.0057
Subject 4	9 times	675	878,937	0.02
Subject 5	3 times	9,346	828,400	0.72

3.2 Non-training phase

Fig. 7 shows the score transition for each of the trials in the non-training set. Table 4 shows the results of applying a linear approximation to each trial and score for the non-training set.

In Fig. 7, it was expected that scores would gradually increase as training continued, and then gradually decrease as training was paused, but this was not the case. It was found that there were variations in scores, such as a decrease in the first training session but an increase in the second training session. In addition, scores in the 700,000 range were seen at the beginning of training, but during this phase, all subjects scored above 800,000 and did not drop below that level.

In the slope of the linear approximation equation, Subject 1 to Subject 3 tended to drop slightly. Subject 5 is not included in the analysis at this time because the number of trials is still small (3) and it is necessary to increase the number of trials in order to compare the data. Subjects 2 and 3 had the same training phase, but subject 2's score decreased more, and the absolute value of the slope was larger than that of subject 1. The scatter of scores is also larger for subject 2.

3.3 Relationship between Rush and Cool and Score during the training phase

Fig. 8 show the score transition and the breakdown of judgment (percentage of Rush, Cool, and Miss) for each subject during the training phase. First, we compare three of the five subjects, Subject 1 to Subject 3, whose training phases were short and whose slopes in Table 4 were negative.

Subject 1's score did not increase until the seventh day, but increased after the eighth day, approaching 900,000 points. During this phase, Rush and Cool were reduced and the Miss rate, in particular, was reduced to 1.7%. Since then, the Miss rate remained low, and was 0% in four instances. When timing is judged as Miss, the score distribution is 0%, resulting in an increase in the Just rate and a significant increase in the score.

Subject 2 continued to increase his score steadily from the second day, reaching a score of 900,000 points on the sixth day. Both Rush and Cool were gradually decreasing, but the Miss rate was unstable, causing the score to decrease over several days.

Subject 3 had a high Cool rate until the third day, but it decreased after the fourth day, and exceeded 900,000 points on the eighth day. Compared to Subject 2, the Miss rate was stable and remained below 1% after the seventh day.

Something that these three subjects have in common is that the rate of Cool is higher than that of Rush on the first day. During the course of the increase in score, there were days when the ratio of Cool to Rush was reversed. This may be due to the fact that the players are not accustomed to playing music games on the first day, so their recognition of the notes flowing from above is not up to par, and their timing may fall a little behind that of Just. Then, it is thought that the sense of rhythm acquired from training experience when the player has become somewhat accustomed to the game will be out of sync with the sense of recognition of the notes, resulting in more Rushes.

Subjects 2 and 3 had the same training phase of 10 days, but the slope in Table 4 is more negative for Subject 2, and there is more variability in the scores. One possible reason for this is the instability of the Miss rate. Subject 2, whose Miss rate was unstable during the training phase, had an average Miss rate of 3.3%, and no Miss rate lower than 1%, even during the non-training phase. Subject 3 maintained a low Miss rate, averaging 0.9%. The percentages of Rush and Cool were lower in Subject 2 , but the difference in Miss rate was larger than that, and the score was judged to be low.

Fig. 8-Sub.4 shows the scores and breakdown of judgments for subject 4. Subject 4's score did not increase and remained stagnant until the 23rd day. However, after that, Rush, Cool, and Miss gradually decreased, and the score reached 900,000 points on the 30th day. In Fig. 7, the score of subject 4 is the most stable, and the values of Rush, Cool, and Miss for subject 4 are also stable with respect to the score just before the end of the training phase.

Subject 5 had the longest training phase among all the subjects, but his score stopped growing around 860,000 points. In Fig. 8-Sub.5, the number of Misses is gradually decreasing, but Rush and Cool are quite unstable. The sum of Rush and Cool averages 25%, only once falling below 20%, and it does not decrease significantly, through to the end of the training phase. In this case, the length of the training phase is not proportional to the increase in score. Rather, the length of the training phase may have decreased the motivation to train, leading to stagnation and instability in scores. The possibility of such a causal relationship is a subject for further investigation.

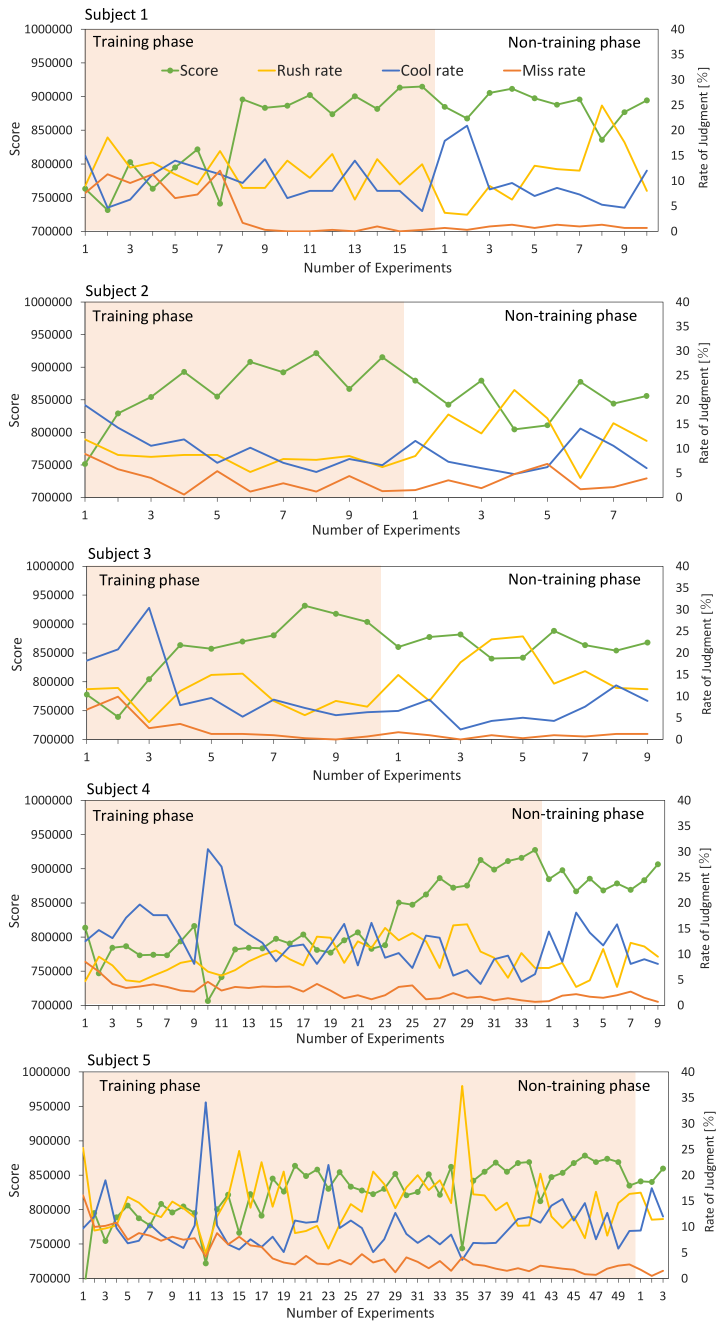

Fig. 8. Transitions of score and rate of judgments

4 Analysis

4.1 Pattern classification for skill loss

Based on the changes in the scores of the five subjects and the breakdown of their judgments, we classified the patterns of skill loss into the following three major categories: a pattern in which the subject forgets gradually after maintaining the score for a while, a pattern in which the score fluctuates wildly, and a pattern in which the subject does not easily forget. For Subject 5, the interval between experiments was irregular, and the frequency of experiments was low, so it was not possible to observe the daily fluctuation of the score. Therefore, we did not classify it as any of the patterns in this study.

Pattern of maintaining for a while and then gradually losing the skill: Subject 1 maintained a high score of around 900,000 points with a slight steady increase until the seventh experiment after entering the phase of training suspension. From the eighth experiment onward, the score exhibited a gradual downward trend (Fig. 7). During this phase, the score slightly decreased during the second experiment, slightly increased during the ninth experiment, and significantly decreased during the tenth experiment, but this is considered to be within the range where it can be called an exceptional phenomenon. We will hereafter continue to examine the trends and correlations in the breakdown of score judgments (ratio of Rush, Cool, and Miss).

Pattern of wildly fluctuating scores: The scores of Subjects 2 and 3 showed relatively large and repeated ups and downs (Fig. 7). The reason for the larger range of fluctuation in Subject 2's score than in Subject 3's score may be due to the instability of the Miss rate during the training phase (Section 3.3). Note that both Subjects 2 and 3 trained for a relatively short phase of time (10 days).

Pattern of not easily losing the skill: Fig. 7 shows that subject 4's score was the most stable and therefore that this subject exhibits a pattern of not easily losing the skill. Subject 4's score and percentage of Rush, Cool, and Miss grew steadily in the second half of the training phase. Empirically, we feel that skills that accumulate steadily during the training phase are less likely to be forgotten during phases of inactivity, and Subject 4 seems to fall into this pattern.

Subjects 3 and 4 have similar training phases, but different skill-loss patterns. First, let us examine the similarities between these subjects. Subjects 3 and 4 are similar in that Cool increases rapidly in the first half of the training phase, after which the ratio of Rush and Cool repeatedly reverses. Another thing these subjects have in common is that their scores and the values of Rush, Cool, and Miss are relatively stable just before the end of the training phase. On the other hand, in terms of the pattern of skill loss, Subject 3's score fluctuates between 800,000 and 900,000, while Subject 4's score remains stable above 850,000, and even exhibits an upward trend after 50 days. If the learning of a skill involves a cognitive process of retention, then the fact that Subject 3 had a short training phase of 10 days may mean that there was insufficient time for the acquired skill to take root.

4.2 Relation between subjects' introspection and scores

Open-ended interviews were conducted with subjects about their play during the training phase inactivity, and subjects were asked to talk about the relationship between their introspection and their scores, as well as their feelings about their play. The overall trend was that subjects who felt they were losing their skill did not experience a decrease in score, while those who did not feel that they were losing the skill did experience this. As for individual comments, these included: "My fingers remember the movements, and my score does not increase at all even if I stop practicing, but when I play after a long time, I find that I cannot complete the parts that used to be easy"; "My score has started to drop because I play only once a week"; "I do not really feel that I am forgetting. Once they are able to do well on that piece, they may be able to maintain a certain score on an easy piece with ease."

5 Conclusion

In this study, we analyzed and classified skill-loss patterns as a preparatory step for constructing a model for predicting the loss of acquired skills in music games. We examined the extent to which subjects forgot, after stopping training, the acquired skill of attaining a certain score for an assigned piece in a music game, then we investigated the relationship between subjects' skill-loss patterns and training patterns. As a result, three types of skill-loss patterns were extracted.

Future work includes investigating whether these skill-loss patterns are applicable to other people and whether they can be generalized. For this purpose, we will increase the number of subjects and continue the experiment to confirm what kind of skill-loss patterns exist.

Subjects 3 and 4 had similar numbers of Rush, Cool, and Miss during the training phase, but their skill-loss patterns were classified differently. To clarify this difference, it is necessary to investigate the relationship between the length of the training phase and the skill-loss patterns. We will also investigate the relationship between Subject 5's training phase duration and score stagnation, as well as the relationship between length of training phase and decrease in motivation. As a future prospect, we would like to improve the content of experiments, for example by altering the time of the experiment, and investigating whether a subject's condition on that day affects the score, and where and how the subject made mistakes during the play.

Acknowledgements

This work was supported by JSPS KAKENHI Grant number 21K18518 and 22H01047.

References

1. Sadao Kawamura, Mizuto Ida, Takahiro Wada, and Jing-Long Wu. Development of a virtual sports machine using a wire drive system-a trial of virtual tennis. In *Proceedings 1995 IEEE/RSJ International Conference on Intelligent Robots and Systems. Human Robot Interaction and Cooperative Robots*, volume 1, pages 111–116 vol.1, 1995.

2. Takuto Nakamura and Hideki Koike. Golf Club-Type Device with Force Feedback for Modifying Club Posture. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI EA '20, page 1–7, New York, NY, USA, 2020. Association for Computing Machinery.
3. Chen-Chieh Liao, Hideki Koike, and Takuto Nakamura. Realtime center of mass adjustment via weight switching device inside a golf putter. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI EA '20, page 1–8, New York, NY, USA, 2020. Association for Computing Machinery.
4. Takuto Nakamura, Daichi Saito, Erwin Wu, and Hideki Koike. Actuated club: Modification of golf-club posture with force feedback and motion prediction in vr environment. In *ACM SIGGRAPH 2020 Emerging Technologies*, SIGGRAPH '20, New York, NY, USA, 2020. Association for Computing Machinery.
5. Kazuyuki Henmi and Tsuneo Yoshikawa. Virtual lesson and its application to virtual calligraphy system. In *Proceedings. 1998 IEEE International Conference on Robotics and Automation (Cat. No.98CH36146)*, volume 2, pages 1275–1280 vol.2, 1998.
6. Saeka Tanaka, Yoshinari Takegawa, and Keiji Hirata. Proposal of a support system for learning brushstrokes in transcription for beginners. In *Proceedings of the 25th International Conference on Intelligent User Interfaces*, IUI '20, page 352–356, New York, NY, USA, 2020. Association for Computing Machinery.
7. Yoshinari Takegawa, Masahiko Tsukamoto, and Tsutomu Terada. Design and implementation of a piano practice support system using a real-time fingering recognition technique. In *Proceedings. 2011 International Computer Music Conference*, pages 1–8, 01 2011.
8. Tomoyasu Nakano, Masataka Goto, and Yuzuru Hiraga. Mirusinger: A singing skill visualization interface using real-time feedback and music cd recordings as referential data. In *Ninth IEEE International Symposium on Multimedia Workshops (ISMW 2007)*, pages 75–76, 2007.
9. Hermann Ebbinghaus. *Memory : a contribution to experimental psychology*. Dover Publications, 1987.
10. Michael W. Eysenck. *Attention and Arousal : Cognition and Performance*. Springer Berlin, Heidelberg, 2012.
11. Charan Ranganath and Robert S. Blumenfeld. Doubts about double dissociations between short- and long-term memory. *Trends in cognitive sciences*, 9(8):374–380, 2005.
12. Cristina M. Alberini. Mechanisms of memory stabilization: are consolidation and reconsolidation similar or distinct processes? *Trends in Neurosciences*, 28(1):51–56, 2005.
13. George Mandler. Organization and memory. *Psychology of Learning and Motivation*, 1:327–372, 1967.
14. Yuko Morita, Keiko Ogawa, and Sunao Uchida. Napping after complex motor learning enhances juggling performance. *Sleep Science*, 9(2):112–116, April 2016. Funding Information: This work was aided by the Japan Society for the Promotion of Science (JSPS) through the Grant-in-Aid for JSPS Fellows. The authors thank Dr. Thomas Maloney for kind and detailed editing of the manuscript. Publisher Copyright: © 2016 Brazilian Association of Sleep.
15. Masako Tamaki, Tatsuya Matsuoka, Hiroshi Nittono, and Tadao Hori. Fast Sleep Spindle (13-15 Hz) Activity Correlates with Sleep-Dependent Improvement in Visuomotor Performance. *Sleep*, 31(2):204–211, 02 2008.
16. Matthew P. Walker, Tiffany Brakefield, Alexandra Morgan, J.Allan Hobson, and Robert Stickgold. Practice with sleep makes perfect: Sleep-dependent motor skill learning. *Neuron*, 35(1):205–211, 2002.