THE EFFECT OF DIGITAL SHEET MUSIC INTERFACES ON BEGINNERS' AUTONOMOUS PRACTICE, FLOW AND COGNITION

Wen-Hsien Kuo[a], Meng-Cong Zheng[1b]

[a] *Department of Industrial Design, National Taipei University of Technology, Taipei, Taiwan*
hsienxx@gmail.com
[b] *Department of Industrial Design, National Taipei University of Technology, Taipei, Taiwan*
zmcdesign@gmail.com

ABSTRACT

With hardware upgrades and the spread of technology, piano learning has shifted from traditional paper scores to digital interfaces offering dynamic guidance and real-time feedback. However, how information presentation in digital sheet music influences beginners' flow experience and cognitive load during self-practice remains underexplored. This study investigates the effects of interface factors (notation format and animation type) on beginners' practice performance and psychological experience. A mixed-method design was adopted using a 2 (notation format: vertical, horizontal) × 2 (animation format: animated image, hand video) between-subject experiment with 40 novices. Participants practiced seven progressively difficult pieces. Post-test measures included the Flow Experience Scale, Cognitive Load Scale, System Usability Scale, and semi-structured interviews. Results showed three main findings: (1) Performance strategy and flow: Participants using sheet music during playing achieved significantly higher flow and concentration than those relying solely on video (p=.004). (2) Interface interaction: The "vertical notation + hand video (Ab)" combination yielded the best flow (M=4.08) and lowest mental effort (M=3.10), indicating optimized element pairing enhances learning immersion. (3) Subjective preference: Although Ab performed best, participants favored Bb (horizontal notation + hand video), likely due to gaze-switching distance. Interview analysis further revealed expectations for predictive cues and hierarchical practice designs to reduce notation-reading errors. These findings suggest specific combinations of notation and animation formats can enhance flow and reduce cognitive load. The study provides concrete directions for digital sheet music design to lower beginners' learning barriers and support immersive piano practice.

Keywords: *User experience, System Usability, Cognitive Load Theory, Optimal Experience*

[1] Corresponding author. 1, Sec. 3, Zhongxiao E. Rd., Taipei 10608 Taiwan, zmcdesign@gmail.com

1 INTRODUCTION

The acquisition of piano performance is a highly labor-intensive activity that requires long-term commitment to structured "deliberate practice" (Ericsson et al., 1993). For beginners, maintaining stable practice quality is the key to developing piano literacy. However, research indicates that without systematic "formal practice," beginners often face bottlenecks of learning stagnation (Bonneville-Roussy & Bouffard, 2015). With the development of digital hardware and multimedia technology, particularly in the post-pandemic era, score formats have shifted from traditional paper to digital interfaces featuring animated guidance, split screens, and real-time evaluation functions (Wang, 2022; Wu, 2023), providing immediate reference resources for autonomous practice. Although digital tools can enhance motivation, improper presentation of interface information may also trigger cognitive overload and hinder more refined musical expression (Rogers et al., 2014). In learning contexts, users' "flow experience" and "cognitive load" are key indicators for measuring learning effectiveness (Engeser & Schiepe-Tiska, 2012; Sweller, 1998). Flow refers to the state of immersion produced when an individual is fully concentrated on a certain activity, which can enhance engagement and a sense of control (Csikszentmihalyi, 2014). On the other hand, imperfect instructional design generates "extraneous cognitive load," consuming the learner's limited working memory capacity (Sweller, 1998). Although digital media aims to provide motivating interfaces, excessive redundant information may lead to cognitive overload in beginners(Klepsch et al., 2017).

Current digital score apps (such as Flowkey, Skoove, etc.) employ diverse visual cues, such as dynamic guidance, fingering video demonstrations, and color-coded prompts (Molero et al., 2021; Wu, 2023). However, the impact of different interface presentation forms—such as the vertical or horizontal layout of the score, and the visualization differences in keyboard dynamic images (animated images versus hand recordings)—on beginners' score-reading behavior and psychological perceptions remains inconclusive. This study aims to investigate the specific effects of different digital interface forms on beginners' practice effectiveness, flow state, and learning cognition.

To achieve the aforementioned objectives, this study employs an experimental design: a 2 (Score Format: Vertical/Horizontal) x 2 (Animation Form: Digital Animation/Hand Recording) two-factor design to evaluate beginners' experiences with different interface combinations. By analyzing the results, interface optimization design and verification will be conducted, aiming to provide specific reference standards for future digital score design.
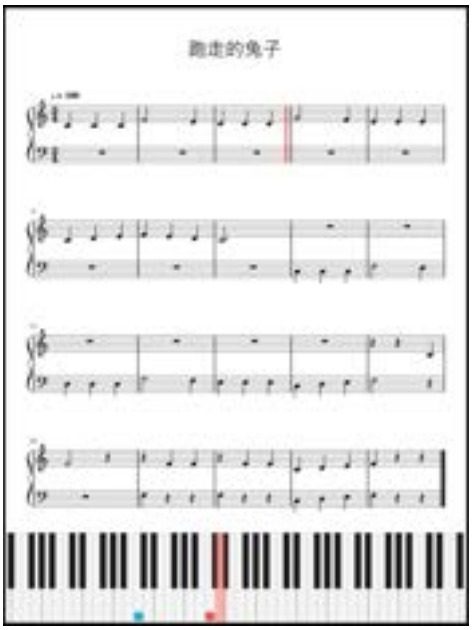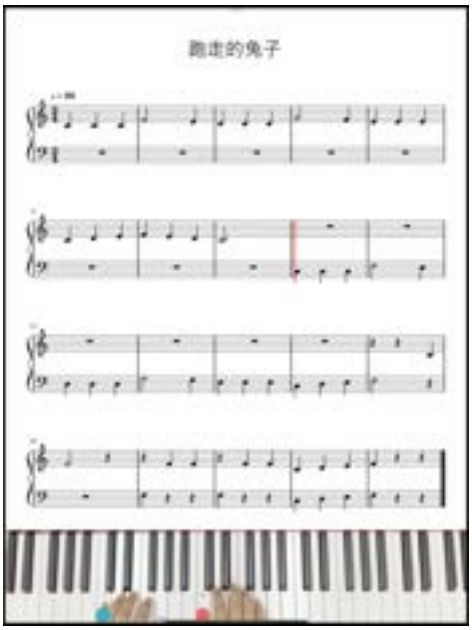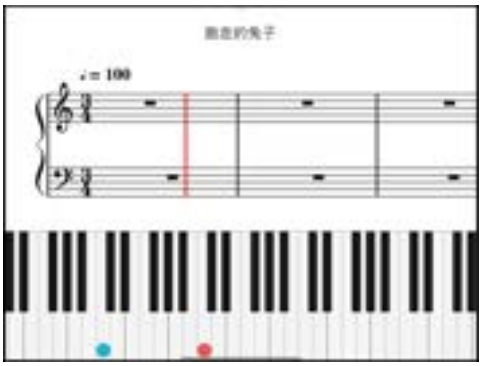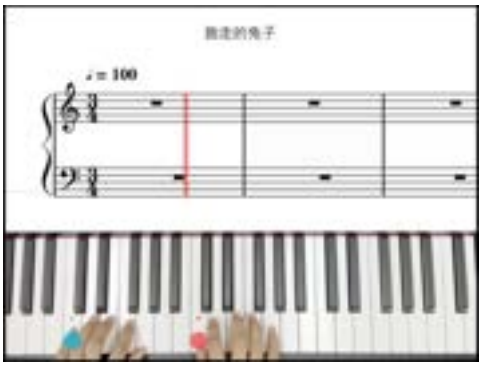
2 METHOD

This study adopts a Mixed Method design for the experiment: by analyzing existing digital musical score software on the market, key design factors affecting learning are clarified.

2.1 Experiment on Digital Musical Score Practice Simulation

The study adopts a 2x2 between-subjects factorial experimental design (Between-subjects Factorial Design). Independent variables are defined as follows: Variable 1: Musical Score Presentation Format (A) Vertical Score: Simulating traditional paper layout, presenting 3 to 4 lines of musical score simultaneously. (B) Horizontal Score: Adopting a single-line horizontal scrolling presentation. Variable 2: Instructional Animation Format (a) Dynamic Image: Using synchronized digital color blocks and virtual keyboard prompts. (b) Hand Recording: Providing performance videos of a real pianists, supplemented with color-coded fingering. Based on the above variables, four sets of experimental samples were combined: Aa (Vertical + Animation), Ab (Vertical + Recording), Ba (Horizontal + Animation), and Bb (Horizontal + Recording). All experimental samples were developed using Adobe After Effects for simulation to ensure precise synchronization of audio, score scrolling, and fingering images.

Table 1. Visual Presentation of Experimental Samples

	
Aa: Vertical Score / Animated Graphics	**Ab**: Vertical Score / Hand Recording
	
Ba: Horizontal Score / Animated Graphics	**Bb**: Horizontal Score / Hand Recording

This experiment recruited 40 participants with completely no piano learning experience (equally distributed into four groups, n=10 per group), to eliminate interference from prior knowledge. Experimental equipment utilized a Yamaha P-525 digital piano and an iPad Pro 12.9-inch screen as the display interface. The experimental tasks were designed in collaboration with professional piano teachers, selecting 7 musical pieces with a difficulty gradient, including 3 right-hand fingering, 3 left-hand fingering, and 1 two-hand ensemble task (such as "Mary Had a Little Lamb," "The Runaway Rabbit," etc.) The procedure was as follows: (1) Preliminary preparation: Conduct 3 to 5 minutes of basic music theory explanation (such as note durations and finger-to-key mapping), and basic scale practice to construct preliminary concepts; (2) Formal experiment: Each musical piece was practiced in sequence with "slow speed (80 BPM) twice" and "normal speed (100 BPM) once," without providing human guidance during the process. After completing the tasks, participants were required to complete the following scales:

1. Flow Experience Scale: Assessing challenge balance, concentration, and immersion (Kiili, 2006; Hou & Chou, 2012).
2. Cognitive Load Scale: Measuring mental load and mental effort levels (Chu et al., 2010; Hwang et al., 2013).
3. System Usability Scale (SUS): Assessing the operability of the interface. Finally, participants were required to perform preference ranking for the four group samples and undergo a semi-structured interview to deeply explore their gaze-switching behavior and reasons for visual dependence (Bangor, 2009).

3 RESULT

This study conducted digital musical score tests of 4 samples on 40 participants without piano learning experience through experiments, and performed analysis from four aspects: subject characteristics, playing strategies, quantitative scales, and interview feedback, to clarify the impact of interface combination forms on learner experience.

3.1 Analysis of Participant Characteristics and Playing Strategies

There were a total of 40 experimental participants (17 males, 23 females), with ages distributed between 18 and 29 years old, and all had no piano learning experience. Analysis of participant performance across seven musical pieces revealed that even as beginners, they exhibited significant differences in "playing strategy differences" when facing playing tasks. Based on behavioral self-assessments, participants were categorized into two primary groups: (1) Non-score using group (50%): Completely relied on the keyboard animation below (100% video), and did not pay attention to the information on the staff. (2) Score using group (50%): Referred to the blocks of the staff to varying degrees during the performance process (30-70% score), attempting to establish a correspondence between fingering and scales.

3.2 Quantitative Impact of Performance Strategies on Learning Experience

To investigate the impact of whether beginners possess "performance strategies" on the learning experience, this study utilized a One-Way Analysis of Variance (One-Way ANOVA) for

testing: Results showed that participants who used scores were significantly higher in total flow scores than those who did not use scores (p=.004). This difference was simultaneously reflected in the two sub-constructs of "flow antecedents" (p=.040) and "flow experience" (p=.006). The data illustrate that when learners actively allocate attention to the musical score, they can establish clearer learning goals and feedback mechanisms, thereby generating high concentration and a positive sense of immersion.

Table 2. One-Way ANOVA of Flow Experience by Performance Strategy

Measurement Variables	Score-Using Group (n=20) *M (SD)*	Non-Score-Using Group (n=20) *M (SD)*	*F*	*p*
Flow Antecedents	3.50 (0.54)	3.13 (0.55)	4.532	.040*
Flow Experience	3.97 (0.52)	3.50 (0.49)	8.57	.006**
Total Flow Score	3.75 (0.42)	3.33 (0.44)	9.51	.004**

* p < .05, ** p < .01

The computational results for cognitive load indicated that there were no significant differences between the two groups in mental load, mental effort, or total cognitive load scores. However, in terms of mean trends, the group without score-reading habits (M = 4.37) was slightly higher than the group with habits (M = 3.71), indicating that not looking at the score may lead to higher mental load. In terms of System Usability Scale (SUS), the evaluations for the two groups were 74 points and 69 points respectively, both reaching the "Good" level, confirming that the interface design possesses sufficient versatility and that the operational threshold will not change significantly due to score-reading behavior.

3.3 Interaction Analysis of Interface Combination Forms

In addition to performance strategies, this study also conducted a specific analysis based on the 2 (Score Format: Vertical/Horizontal) x 2 (Animation Style: Animated Graphics/Hand Recording) two-factor experimental design, as follows:

Through Two-Way Analysis of Variance (Two-Way ANOVA), it was found that "score format" and "animation format" have significant interactions in flow experience (p = .020), total flow score (p = .019) (Table 3), and cognitive mental effort (p = .009) (Table 4). Data show that the "vertical score + hand recording (Ab)" combination can create the highest flow experience (M = 4.08) and reduce mental effort to the minimum (M = 3.10). The "horizontal score + hand recording (Bb)" was relatively poor in its performance in these two items (Flow experience: M = 3.46, mental effort: M = 3.97). This indicates that when paired with hand recording (b), the vertical score (A) is the best combination.

3.4 Subjective Preferences and Interview Findings

At the end of the experiment, participants were invited to rank the preferences of the four samples and undergo qualitative interviews to understand the operational context behind the participants. According to the Friedman test, the highest subjective preference among all participants was Bb (horizontal score + hand recording) (p < .001). However, if subdivided by strategy groups, the "Score Group" most preferred Ab (M=1.95), while the "No-score Group"

most preferred Bb (M=1.55). This result illustrates that the realism provided by hand recording is attractive to all beginners, but in terms of score layout, those with score-using strategies are more inclined toward a vertical layout.

Table 3. Two-Way ANOVA of Flow Experience Across Four Samples

Dependent Variable	Source of Variation	SS	df	MS	F	p	ηp2
Flow Antecedents	Score Format	0.01	1	0.01	0.04	.845	.001
	Animation Style	0.18	1	0.18	0.57	.454	.016
	Interaction	0.87	1	0.87	2.74	.107	.071
	Error	11.44	36	0.32			
Flow Experience	Score Format	0.47	1	0.47	1.74	.196	.046
	Animation Style	0.03	1	0.03	0.13	.725	.003
	Interaction	1.6	1	1.6	5.92	0.02*	.141
	Error	9.74	36	0.27			
Total Score	Score Format	0.1	1	0.1	0.51	.480	.014
	Animation Style	0.09	1	0.09	0.42	.520	.012
	Interaction	1.24	1	1.24	6.04	0.019*	.144
	Error	7.39	36	0.21			

* p < .05, ** p < .01

Table 4. Two-Way ANOVA of Cognitive Load Across Four Samples

Dependent Variable	Source of Variation	SS	df	MS	F	p	ηp2
Mental Load	Score Format	0.2	1	0.2	0.12	.730	.003
	Animation Style	1.3	1	1.3	0.8	.378	.022
	Interaction	2.12	1	2.12	1.3	.261	.035
	Error	58.46	36	1.62			
Mental Effort	Score Format	0.03	1	0.03	0.02	.881	.001
	Animation Style	1.47	1	1.47	1.34	.256	.036
	Interaction	8.4	1	8.4	7.63	.009**	.175
	Error	39.63	36	1.1			
Total Score	Score Format	0.11	1	0.11	0.09	.767	.002
	Animation Style	1.36	1	1.36	1.07	.307	.029
	Interaction	3.98	1	3.98	3.15	.084	.080
	Error	45.57	36	1.27			

* p < .05, ** p < .01

Table 5. Friedman Test for Overall Preference Ranking: Non-parametric Comparison of Multiple Dependent Samples

Sample Combination	*n*	*M*	*SD*	Mean Rank	Preference Rank
Bb	40	1.78	0.97	1.78	1
Aa	40	2.15	0.98	2.15	2
Ba	40	2.78	1	2.78	3
Ab	40	3.3	0.91	3.3	4

Results of the Friedman test：(3) = 32.73,** $p < .001$

Table 6. Friedman Test for Preference Ranking Grouped by Performance Strategy: Non-parametric Comparison of Multiple Dependent Samples

Sample Combination	Score-Using (n=20)			Non-Score-Using (n=20)		
	M	*SD*	Preference Rank	*M*	*SD*	Preference Rank
Ab	1.95	0.95	1	2.35	0.99	2
Bb	2.00	1.12	2	1.55	0.76	1
Ba	3.00	0.92	3	2.55	1.05	3
Aa	3.05	1.05	4	3.55	0.69	4

Friedman test results for the Score-Using Group:$\chi^2(3) = 13.26$, $^{**}p = .004$
Friedman test results for the Non-Score-Using Group:$\chi^2(3) = 24.36$, $^{**}p < .001$

Through the organization of participant interview content, the following core findings were summarized: Beginners exhibit an extreme reliance on interface guidance elements (such as red lines and color blocks) during the practice process and regard real hand recordings as a core tool for alleviating the pressure of scale recognition.

For participants with a score-reading inclination, their learning behavior follows a hierarchical path from imitating fingering movements to establishing sensory melodic memory, and finally to verifying musical staff symbols, indicating that digital interfaces can effectively serve as a transitional bridge before entering the domain of traditional score reading. However, due to the perceptual latency caused by the synchronized display of prompts and music in existing systems, participants are often unable to react in time, leading to a strong demand for visual anticipation, such as displaying the next key position in advance.

Especially when facing high-demand passages involving two-hand coordination or wide-scale movements, the sudden increase in cognitive load frequently caused participants to exceed their processing threshold, resulting in a complete abandonment of score reading and regression to a passive, video-following mode. Notably, melodic familiarity played a significant moderating role in this process: prior memory of the melody effectively reduced visual dependence on the interface, transforming the system's function from a dominant navigation role into a supportive reference tool.

4 DISCUSSION

This study explores the impact of digital music score interfaces on beginners' playing practice through experiments and analysis. The following is a deep discussion on playing strategies, interface fitness, and the gap between subjective and objective evaluation.

4.1 The Key Role of Playing Strategies on Learning Immersion

Research results show that whether participants adopt the strategy of "reading music scores" significantly affects their perceived level of flow. Although it is highly challenging for those without piano learning experience to read scores and play simultaneously, statistical data shows that the score-using group outperformed the non-score-using group in total flow score (p=.004), flow antecedents (p=.040), and flow experience (p=.006). This echoes Csikszentmihalyi's definition of flow antecedents: clear goals and immediate feedback can guide the focus of attention (Csikszentmihalyi, 2014). When beginners attempt to establish logical connections between scales and fingerings, music scores provide structured task guidance, which in turn transforms into higher concentration and positive immersion.

4.2 Interaction of Interface Elements and Patterns of Fitness

This study confirmed through two-way ANOVA that the quality of digital score interfaces is not determined by a single element, but depends on the interactive pairing of score format and animation format. Best Performance Combination (Ab): Experimental data show that "vertical score + hand recording" performed best in enhancing flow experience (M=4.08) and reducing cognitive mental effort (M=3.10).

The interview results further confirm that the vertical score can provide better information continuity and preview space, while hand recording can effectively reduce the conversion pressure between score scales and actual piano key mapping, reducing the burden of gaze-switching. The research results also point out significant compatibility differences between interface elements; specifically, the vertical score is suitable for pairing with concrete hand recording, while the horizontal score is more compatible with geometric animated graphics. This illustrates that under different visual layouts, there are differences in the brain's efficiency in processing concrete actions and abstract symbols; therefore, future design adjustments should deeply consider the gaze-switching distance and the complexity of information processing.

4.3 Analysis of the Gap Between Subjective Preference and Objective Performance

Another noteworthy finding is the gap between participants' "preference" and "performance." Although the Ab combination performed best in flow and cognitive load indicators, according to the Friedman test, the favorite of all participants was Bb (horizontal score + hand recording). This may be because the horizontal layout makes the proportions of the score and animation blocks similar, with short visual movement distances, giving beginners an intuitive feeling of "easier concentration." However, because music and images appear simultaneously, even with high focus on the screen, delays still occur, causing participants to fall into a state of passive following.

4.4 Optimization Suggestions for Digital Music Scores

Combining the results of quantitative data and qualitative interviews, this study proposes the following optimization directions for digital sheet music interfaces. First, to alleviate the perceptual latency and reaction discrepancy caused by synchronized display, a visual anticipation mechanism should be introduced to pre-announce subsequent note positions in advance, thereby reducing the reaction pressure on participants.

Second, according to the participants' three-stage behavioral strategy of "imitation, memory, and verification," a hierarchical practice mode should be established. For example, fingering video guidance should be strengthened in the initial stage to establish muscle memory, while video navigation should be gradually weakened and staff symbols strengthened in the later stages to assist users in transitioning from system dependence to the goal of independent score-reading and performance.

Finally, to reduce the cognitive dissonance and visual fatigue caused by multiple dynamic elements, the interface layout should be optimized to reduce dynamic interference. Especially when performing high-difficulty tasks such as hands-together ensembles, background information should be simplified or clearer early-warning mechanisms should be provided to avoid the psychological frustration caused by rhythmic breakdown in beginners.

5 CONCLUSION

This study reveals the profound impact of digital music score interface design on piano beginners' autonomous practice. The research confirms that guiding users to actively read scores during the practice process, rather than only passively imitating video prompts, can significantly enhance immersion and concentration levels during learning. In addition, a gap exists between beginners' subjective preferences and actual operational performance, illustrating that interface design should transcend simple visual intuition and prioritize resolving obstacles such as perceptual delay that affect rhythmic accuracy. Future design optimization should introduce guidance functions with "anticipation mechanisms" and combine them with the "imitation—memory—verification" hierarchical teaching path. Through this progressive auxiliary framework, learners can be assisted in transitioning smoothly from complete reliance on navigation to independent performance, achieving a more ideal autonomous learning experience.

REFERENCE

Bonneville-Roussy, A., & Bouffard, T. (2015). When quantity is not enough: Disentangling the roles of practice time, self-regulation and deliberate practice in musical achievement. *Psychology of Music*, *43*(5), 686–704. https://doi.org/10.1177/0305735614534910

Csikszentmihalyi, M. (2014). *Applications of Flow in Human Development and Education: The Collected Works of Mihaly Csikszentmihalyi*. Springer Netherlands. https://doi.org/10.1007/978-94-017-9094-9

Engeser, S., & Schiepe-Tiska, A. (2012). Historical Lines and an Overview of Current Research on Flow. In S. Engeser (Ed.), *Advances in Flow Research* (pp. 1–22). Springer New York. https://doi.org/10.1007/978-1-4614-2359-1_1

Ericsson, K. A., Krampe, R. T., & Tesch-Romer, C. (1993). *The Role of Deliberate Practice in the Acquisition of Expert Performance*.

Klepsch, M., Schmitz, F., & Seufert, T. (2017). Development and Validation of Two Instruments Measuring Intrinsic, Extraneous, and Germane Cognitive Load. *Frontiers in Psychology*, *8*, 1997. https://doi.org/10.3389/fpsyg.2017.01997

Molero, D., Schez-Sobrino, S., Vallejo, D., Glez-Morcillo, C., & Albusac, J. (2021). A novel approach to learning music and piano based on mixed reality and gamification. *Multimedia Tools and Applications*, *80*(1), 165–186. https://doi.org/10.1007/s11042-020-09678-9

Rogers, K., Röhlig, A., Weing, M., Gugenheimer, J., Könings, B., Klepsch, M., Schaub, F., Rukzio, E., Seufert, T., & Weber, M. (2014). P.I.A.N.O.: Faster Piano Learning with Interactive Projection. *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces*, 149–158. https://doi.org/10.1145/2669485.2669514

Sweller, J. (1998). *Cognitive Architecture and Instructional Design*.

Wang, Y. (2022). Creative Digital Tools for Teaching Piano. *Revista Música Hodie*, *22*. https://doi.org/10.5216/mh.v22.70961

Wu H.-Y. (2023). *The Effects of Incorporating Tactile Prompts into a Piano Fingering Practice Application on Flow Experience, Cognitive Load, and Learning Emotions for Piano Beginners*.