

Journal of Digital Education and Learning Engineering

ดำเนินการวารสารโดย สมาคมการศึกษาดิจิทัลและวิศวกรรมการเรียนรู้

Designing and Evaluating PinyinDash: A Dual-Mode Digital Game Prototype for Chinese Character-Pinyin Learning among Thai High School Learners

Zitong Zhang¹, Kannika Daungcharone^{2*} & Kittawaya Thongkoo²

Multidisciplinary and Interdisciplinary School, Chiang Mai University, Thailand¹, College of Arts, Media and Technology,
Chiang Mai University, Thailand²

Received: June 14, 2026 Revised: June 21, 2026 Accepted: June 24, 2026

Abstract

Chinese character-pinyin mapping is a foundational but demanding task for beginning learners of Chinese as a foreign language. Learners must connect visual character forms with alphabetic pinyin spelling and tonal information, yet classroom practice often relies on copying, word lists, or recognition-based exercises that provide limited active retrieval and feedback. This study presents PinyinDash, a dual-mode digital game prototype designed to support Chinese character-pinyin learning through accurate pinyin-tone input, post-learning review, rapid memory retrieval under time pressure, light pre-exposure, immediate feedback, and customizable classroom word banks. The prototype combines Learning Mode, which organizes characters into short tone-marked pinyin input sessions, and Time Mode, which turns toneless pinyin retrieval into a post-learning falling-meteor review activity. Time Mode outcomes are interpreted as retrieval-practice and pre-exposure signals rather than direct evidence of tone-marked mastery. The design was informed by digital game-based learning, retrieval practice, spaced review, and Keller's Attention, Relevance, Confidence, and Satisfaction (ARCS) motivation model. A one-group pre-test/post-test classroom study was conducted with 30 Grade 10 students in Chiang Mai, Thailand. After teacher instruction on the target item pool, students completed a pre-test, used PinyinDash for 2 weeks, and completed a post-test and an ARCS-based motivation questionnaire. Learning achievement increased from the pre- test ($M = 13.00$, $SD = 3.19$) to post-test ($M = 18.37$, $SD = 2.19$), $p < .001$. Motivation results showed high agree/strongly agree response percentages across Relevance (95.0%), Satisfaction (92.5%), Attention (88.3%), and Confidence (85.8%), with high overall internal consistency (Cronbach's $\alpha=.928$). These findings suggest that PinyinDash is a feasible and motivationally acceptable prototype for classroom Chinese character-pinyin practice, while controlled studies are needed to establish causal effectiveness.

*Corresponding author

Email address: kannika.d@cmu.ac.th

Keywords: Chinese language, pinyin, digital game-based learning, ARCS model, learning achievement, prototype evaluation

■ Introduction

Chinese character-pinyin mapping is one of the first technical barriers encountered by learners of Chinese as a foreign language. To participate in even basic classroom tasks, learners must connect a visual character with its alphabetic pinyin spelling, tone, and pronunciation. This mapping is not transparent to beginners because the written character does not directly reveal pronunciation, as alphabetic writing systems often do. Learners, therefore, need repeated encounters with the same forms, but repetition alone is not enough. They also need active recall, feedback, and opportunities to notice and correct partial errors in spelling or tone. In many classroom settings, pinyin and character practice still depend heavily on word lists, copying, teacher-led questioning, or recognition-based tasks. These activities can be useful for initial exposure, but they may not provide enough retrieval practice. A learner who can recognize a correct pinyin form in a multiple-choice item may still struggle to produce that form independently. Similarly, a learner who can copy pinyin with tone marks may not yet retrieve it fluently when a character appears without support. This gap between recognition and production is especially relevant for early-stage learners who are trying to build reliable form-sound links.

Digital game-based learning is one way to address this problem because it can turn repeated practice into a goal-directed, feedback-rich activity. Prior work has suggested that language-learning games can support engagement, participation, and vocabulary practice when game mechanics align with learning tasks rather than serve as superficial rewards (Butler, 2015; Chen & Hsu, 2020; Zhonggen, 2018). Research on vocabulary and game-based tasks has also reported positive or supportive effects on vocabulary achievement, motivation, self-confidence, and word knowledge across different language learning contexts (Li, 2021; Mokhtar & Rahelan, 2024; Rasti-Behbahani & Shahbazi, 2020). Within Chinese learning, recent studies have examined gamified vocabulary tools, Mandarin learning games, and expert or learner perspectives on digital games for Chinese vocabulary learning (Rawendy et al., 2017; Ying et al., 2021; Phupunna & Hongsamsibkao, 2024; Zhang & Abdul Salam, 2025). These studies suggest that game-based tools can be pedagogically useful, but they also point to the need for tools that are designed around specific features of Chinese, including character-pinyin mapping. The need for task-specific design is especially clear in the context of Chinese pinyin learning. A game that only asks learners to select a matching card may increase exposure, but it may not require learners to construct the pinyin form. A game that asks learners to react quickly may increase excitement, but it may also reduce attention to tones if tonal accuracy is not included somewhere else in the learning cycle. For this reason, the present study treats prototype structure as part of the research problem. The question is not only whether students like a game, but whether the game organizes the right kinds of practice: accurate production, repeated retrieval, feedback after errors, and content alignment with what the teacher is actually teaching.

PinyinDash was developed as a targeted response to this need. Rather than presenting Chinese vocabulary solely as a quiz or recognition task, the prototype asks learners to type pinyin. Its core design idea is that beginners need two complementary forms of practice: slow, careful input for accurate pinyin-tone encoding, and post-learning review that encourages rapid memory retrieval after initial study. Therefore, the prototype contains two linked modes. Learning Mode supports accurate tone-marked pinyin input with immediate feedback and retry opportunities, while Time Mode transforms toneless pinyin retrieval into a falling-meteor review game that can also provide light pre-exposure to future learning items. The game also supports teacher-aligned custom word banks, visible progress, score feedback, wrong-item collection, and review support. The present article reports PinyinDash as a research article rather than as a software manual. It focuses on the educational concept, prototype structure, and classroom evidence from a one-group classroom study. The study asks whether this dual-mode prototype is feasible and motivationally acceptable for Thai high school learners of Chinese, and whether learners show improvement in character-pinyin learning after a two-week classroom intervention. Because the empirical design uses a single participating group without a separate comparison group, the study does not provide strong causal evidence. Instead, it provides preliminary classroom evidence to guide future controlled studies and further prototype refinement.

■ Significance and Purposes

- 1) To describe the game concept and prototype structure of PinyinDash in relation to Chinese character-pinyin learning.
- 2) To examine learners' motivation after using PinyinDash through an ARCS-based questionnaire.
- 3) To evaluate learners' achievement change through pre-test and post-test scores from the participating group.

■ Literature Reviews

Digital Game-Based Learning for Language Practice

Digital game-based learning is commonly discussed to promote attention, persistence, interaction, and repeated practice. In language learning, games may be especially useful when they make learners repeatedly process target forms while maintaining a sense of challenge and progress. However, the pedagogical value of a game does not come solely from entertainment. It depends on how well the mechanics of the game are aligned with the learning objective (Butler, 2015; Chen & Hsu, 2020). If learners spend most of their effort on actions unrelated to the target knowledge, game mechanics can distract from learning. If the game action is the learning action, repeated play can become repeated practice.

This alignment is important for vocabulary learning. Word learning usually requires multiple encounters, gradual strengthening of form-meaning or form-sound associations, and feedback after errors. Recent studies on word games and classroom gaming tools have reported supportive patterns for vocabulary development when activities provide active engagement and meaningful repetition (Bustrillo et al., 2024; Mokhtar & Rahelan, 2024). Research on serious games and digital game-based tasks has similarly suggested that interactive vocabulary practice may influence learning outcomes when it is structured around the content to be learned (Chen & Hsu, 2020; Rasti-Behbahani & Shahbazi, 2020; Zhonggen, 2018).

At the same time, a positive game effect should not be assumed simply because an activity is digital or game-like. Game-based practice may help when it increases time on tasks, encourages repeated responses, and provides meaningful feedback, but it may be less useful when points, animation, or competition are disconnected from the learning target. In a vocabulary game, the most important action should involve the target vocabulary itself. This principle guided the design of PinyinDash: scoring, progress, and completion are tied to pinyin responses rather than to unrelated navigation or chance. The prototype, therefore, aims to make the game action and the learning action the same.

In Chinese learning contexts, game-based tools face a more specific design challenge. Learners must connect characters, pinyin spelling, tones, pronunciation, and meaning. Many vocabulary games support recognition or matching, but fewer are built around active pinyin production. Previous work on Chinese language learning games and Mandarin learning tools has illustrated the potential of gamified practice and the importance of connecting game design with vocabulary content, learner engagement, and motivational support (Rawendy et al., 2017; Ying et al., 2021; Phupunna & Hongsamsibkao, 2024; Zhang & Abdul Salam, 2025). PinyinDash addresses this need by treating pinyin input as the core game action rather than a secondary display feature.

This focus on pinyin input also distinguishes the prototype from broader Chinese vocabulary games. Many useful tools combine images, meanings, pronunciation, or character recognition, but beginners often need a narrower and more repetitive bridge between a character prompt and the pinyin form they must produce. A classroom prototype can therefore be valuable even if it does not attempt to teach every aspect of Chinese vocabulary at once. PinyinDash deliberately narrows the task so that learners repeatedly answer the same central question: when this character appears, can I produce the corresponding pinyin form accurately or quickly enough for the present mode?

Retrieval Practice and Spaced Review

Retrieval practice provides a cognitive rationale for input-based character-pinyin practice. When learners actively recall a target form, retrieval itself can strengthen later retention more effectively than passive exposure (Roediger & Karpicke, 2006; Karpicke & Roediger, 2008; Roediger & Butler, 2011). Retrieval-based learning has also been shown to support meaningful learning beyond simple restudy (Karpicke & Blunt, 2011). This is directly relevant to pinyin learning because the learner must retrieve the pinyin spelling from a character prompt. A multiple-choice task may allow recognition, but typing requires the learner to

assemble and produce the form. For beginners, this production requirement can reveal partial knowledge: a learner may remember the initial consonant but forget the final, or may type the correct letters but omit the tone.

Spaced review is also relevant because vocabulary and form-sound associations are not normally stabilized after a single correct response. Distributed practice has been shown to support later retention across many verbal learning tasks (Cepeda et al., 2006), and the timing of spacing can influence later retention (Cepeda et al., 2008). A classroom game cannot fully replace a long-term memory system, but it can provide short-cycle review and prevent difficult items from disappearing too quickly. PinyinDash therefore uses visible mastery states and review support to keep weak, new, and due-for-review items available for continued practice.

For beginning learners, spacing also has a practical classroom meaning. Learners may meet a word during the teacher's explanation, answer it once during an exercise, and then move on before the form is stable. A review-oriented prototype can create additional encounters without requiring the teacher to manually track every learner's weak items. In PinyinDash, review support is intentionally presented through simple visible states rather than through technical terminology. Learners see that some items are new, pre-exposed, in progress, due for review, or mastered. This design keeps the interface understandable while still encouraging distributed practice (Cepeda et al., 2006, 2008).

The distinction between accurate encoding and review-oriented retrieval is central to the present prototype. Learning Mode prioritizes careful pinyin transcription with tone marks. Time Mode prioritizes rapid retrieval of toneless pinyin after learning and can lightly expose learners to items that will be formally studied later. This distinction is pedagogically important because tone-marked accuracy, memory retrieval under time pressure, and prior exposure are related but not identical learning functions. A timed game can support repeated retrieval, but it should not be treated as complete evidence that a learner has mastered pinyin spelling and tone. The design separates these functions while allowing both to contribute to the overall learning experience.

ARCS Motivation Model

Motivation is a major concern in repeated vocabulary practice because the activity can easily become monotonous. Keller's Attention, Relevance, Confidence and Satisfaction (ARCS) model provides a useful framework for designing and evaluating the motivational quality of instructional materials (Keller, 2010). The model includes four dimensions: Attention, Relevance, Confidence, and Satisfaction. Attention concerns whether the learning experience captures and sustains interest. Relevance concerns whether learners perceive the activity as connected to their needs or goals. Confidence concerns whether the environment helps learners believe they can succeed. Satisfaction concerns whether learners experience reward, accomplishment, or positive reinforcement.

The ARCS model is also useful because it separates different motivational qualities that might otherwise be merged into a general statement that learners enjoyed the game. A learner may find the

falling-meteor screen exciting but still feel that the task is not relevant to class. Another learner may believe the content is relevant but feel anxious if the task is too difficult. By reporting Attention, Relevance, Confidence, and Satisfaction separately, the present study can more directly connect learner responses to prototype design decisions. This is important for a design-oriented article because motivation results should inform revision, not only provide a positive evaluation.

PinyinDash used the ARCS model not only as a post-use questionnaire framework, but also as a prototype design lens. In line with Keller's motivational design model, the prototype was designed to address learners' attention, relevance, confidence, and satisfaction during repeated pinyin practice (Keller, 2010). The design also followed the view that game mechanics should be meaningfully aligned with the target learning action rather than function only as external decoration or reward (Butler, 2015; Chen & Hsu, 2020; Zhonggen, 2018). Recent work on Chinese vocabulary educational games has also used the ARCS model to frame game design, supporting its relevance for this context (Zhang & Abdul Salam, 2025). Table 1 summarizes the main ARCS-based design mapping. The table focuses on representative design features rather than listing every interface element.

Table 1.

ARCS-based design mapping

ARCS dimension	PinyinDash features	Design rationale
Attention	Falling Chinese-character meteors that turn repeated pinyin retrieval into a visually changing task; dual time pressure through a 60-second Time Mode round and individual falling-meteor deadlines; immediate visual feedback after correct or incorrect responses.	These features sustain attention during repeated pinyin review by combining visual movement, bounded time pressure, and immediate feedback while keeping the game action tied to pinyin responses.
Relevance	Teacher-aligned custom word banks that allow practice items to match classroom-taught vocabulary; a shared item pool across Learning Mode and Time Mode that connects slow learning and fast review to the same lesson content; light Time Mode pre-exposure that makes some future Learning Mode items familiar before formal study.	These features help learners perceive the game as connected to class learning rather than as unrelated entertainment.
Confidence	Five-item Learning Mode groups that break the word bank into manageable practice units; mistake	Reduces overload, supports step-by-step

ARCS dimension	PinyinDash features	Design rationale
	cards, “Don’t Know” support, and retry opportunities that make errors recoverable; visible mastery states that show each item’s learning status.	success, and makes mistakes part of guided practice rather than final failure.
Satisfaction	Success cards and meteor destruction that make correct retrieval immediately visible; score and combo rewards that turn repeated correct responses into achievement feedback; group-completion review and visible progress indicators that give learners a sense of closure; an optional local leaderboard that records Time Mode achievement.	These features make successful practice visible, rewarding, and complete.

The mapping should be interpreted as design rationale rather than feature-level causal evidence. Several features may support more than one motivational quality, but the table identifies their primary intended role in the prototype. The present study reports learners’ overall motivational response to PinyinDash and cannot determine which individual feature contributed most to the ARCS ratings.

Research Gap

The literature suggests that game-based learning can support vocabulary practice, retrieval practice can strengthen memory, spaced review can organize repeated exposure, and ARCS can guide motivational design. However, there remains a practical need for classroom-oriented prototypes that combine these ideas for the specific problem of Chinese character-pinyin learning. The gap addressed in this study is therefore not simply whether games can be fun, but whether a pinyin-input game can be structured as a coherent learning environment and show promising classroom evidence in both achievement and motivation.

■ PinyinDash Game Concept and Prototype Structure

Overall Design Concept

As shown in Figure 1, PinyinDash was designed as a dual-mode classroom game prototype for Chinese character-pinyin practice. The system uses a teacher-taught item pool or custom word bank as the shared content base. Learning Mode supports slower and more accurate pinyin-tone input, while Time Mode turns later review into a falling-meteor retrieval game. This structure reflects the idea that beginning

learners need both careful form practice and repeated retrieval opportunities, but these functions are better supported through different interaction conditions.

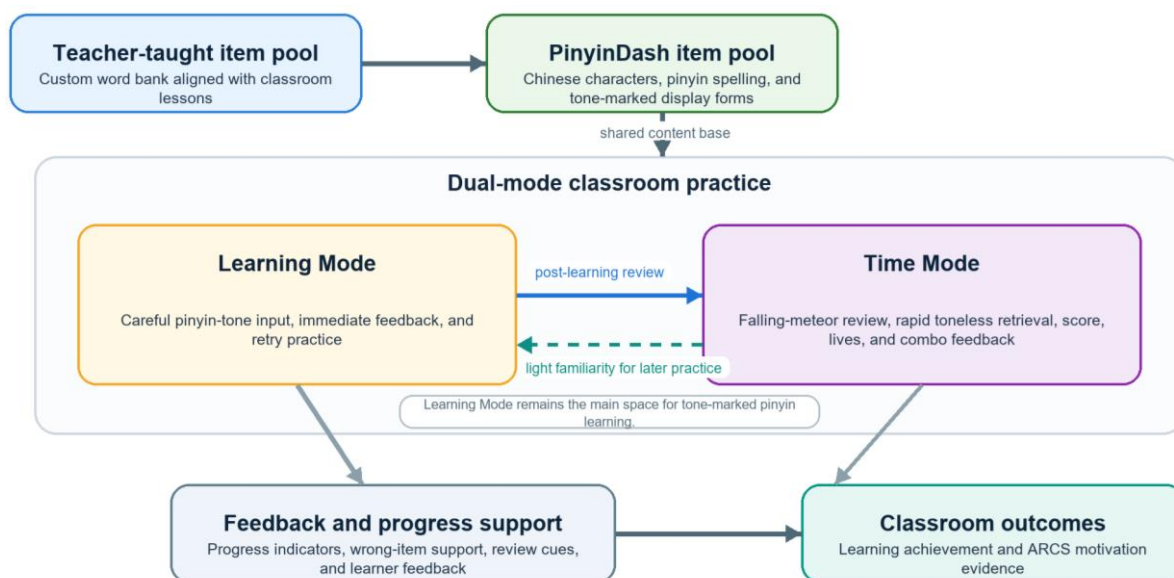


Figure 1. Overview of the PinyinDash classroom practice flow.

In this paper, the game concept is presented at the prototype and classroom-use level. The focus is on how the tool organizes learning activities for students, not on the full internal scheduling mechanism. More technical details of adaptive item selection and scheduling are treated as future system-development work.

Custom Word-Bank Management

Figure 2 illustrates the word-bank management interface, including built-in word banks, teacher-created lists, item viewing, and import functions. A key design feature of PinyinDash is its custom word-bank function. Teachers can prepare or select a word bank containing Chinese characters, pinyin spelling, and tone-marked display forms. The prototype supports both built-in word banks and teacher-created word banks. Teachers may select an existing word bank for general HSK level 1-2 practice or create a classroom-specific list by entering items manually in game or importing a prepared text file. Each item includes a Chinese character or word, a toneless pinyin form for answer matching, and a tone-marked display form for learner feedback. Word banks can be edited, deleted, exported, and imported, allowing teachers to align PinyinDash with lesson content, tests, or review activities. This allows the game content to remain aligned with classroom lessons and assessment content. In the present study, the pre-test, post-test, and game practice were all connected to the same teacher-taught item pool.

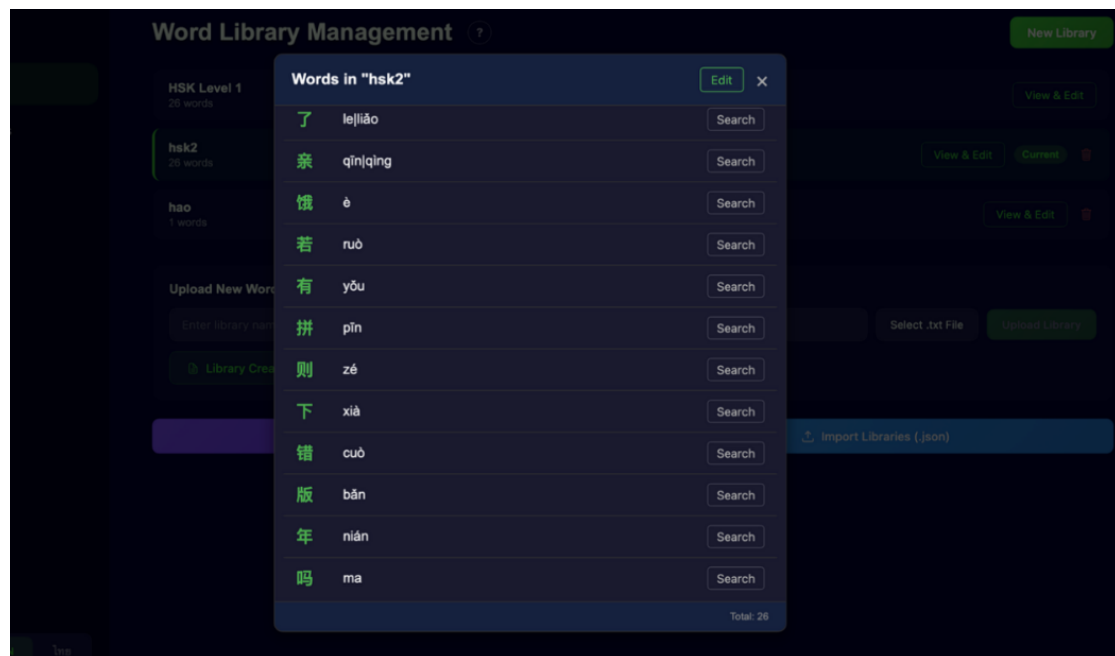


Figure 2. PinyinDash's word bank management interface.

This function supports classroom relevance because students practice items that belong to their current course rather than an unrelated commercial word list. It also preserves the teacher's role in selecting learning content, while the game provides repeated practice, feedback, and review opportunities.

Learning Mode

Figure 3 shows the Learning Mode dashboard, where learners view grouped characters and mastery-related information before practice. Figure 4 shows the pinyin-tone input screen, where learners enter tone-marked pinyin for a target Chinese character. Learning Mode is the accuracy-oriented part of the prototype. It organizes practice into short groups of items so that learners can focus on a manageable set of Chinese characters at one time. During practice, learners see a Chinese character and enter the corresponding pinyin with tone information. The tone-input mechanism differed according to device type. On devices with a physical keyboard, learners typed pinyin letters with the standard keyboard and used the number keys 1-4 to enter the four Mandarin tones. The number key was applied to the most recently typed vowel in the current input sequence. For example, to enter “ǎ”, a learner could type a and then 3, so that the system converted “a” into “ǎ”. On mobile or touch-screen devices, Learning Mode displayed a custom virtual keyboard with letter keys, a backspace key, a confirmation key, and four tone keys corresponding to tones 1-4. This input design allowed learners to produce tone-marked pinyin without installing an external Chinese input method. Immediate feedback helps learners notice spelling or tone errors, and incorrect items can reappear for further practice.

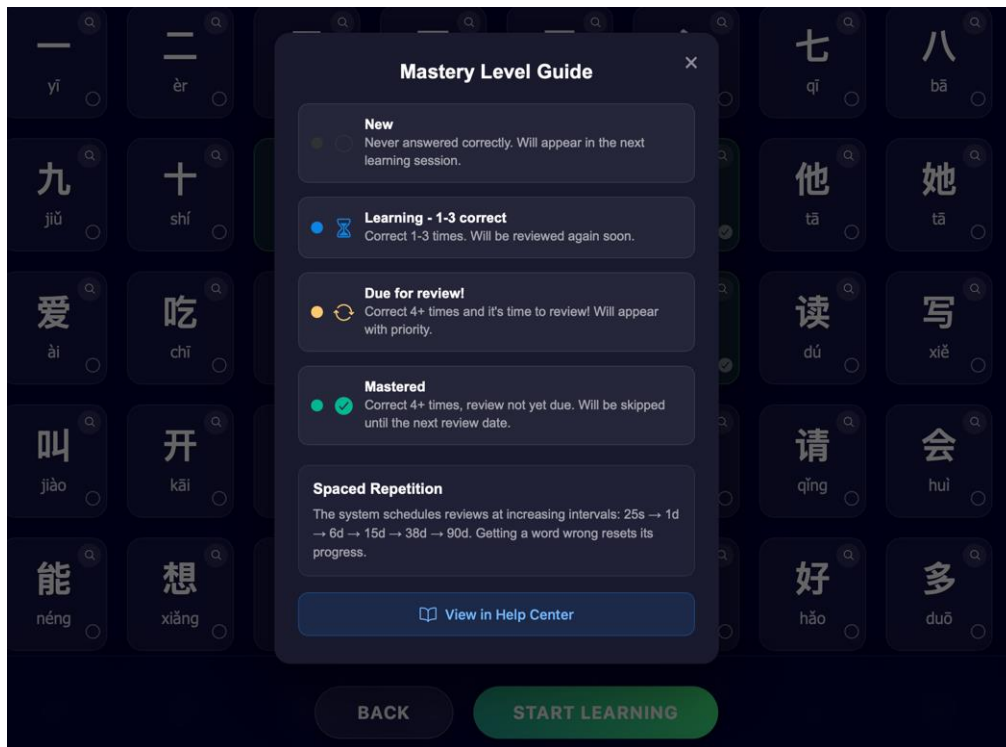


Figure 3. PinyinDash's Learning Mode dashboard.

This mode is important because it requires active pinyin production rather than only recognition. It also keeps tone accuracy in the learning cycle. The input design aims to reduce the technical difficulty of typing tone-marked pinyin while still requiring learners to attend to tones.

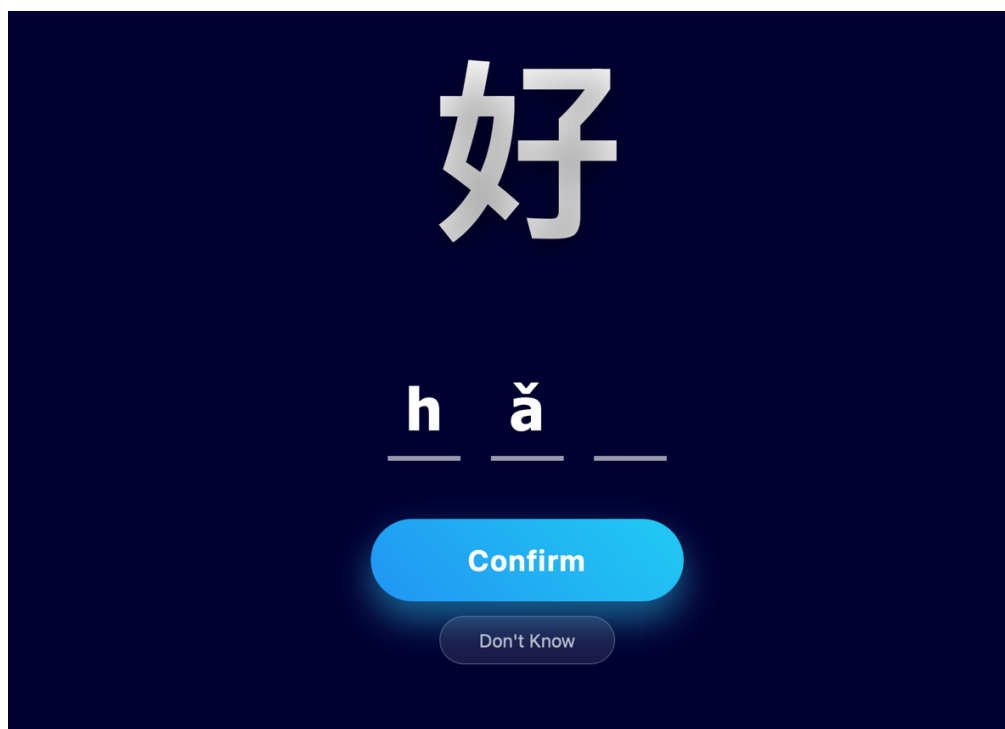


Figure 4. PinyinDash's pinyin-tone input screen.

Time Mode

Figure 5 shows the Time Mode screen, including falling character meteors, score and combo feedback, and the timed retrieval environment. Time Mode is the review-oriented part of the prototype. In this mode, Chinese characters appear as falling meteors, and learners type the corresponding toneless pinyin before the meteor reaches the bottom of the screen. The standard round uses a short time limit, lives, score, and combo feedback to make repeated retrieval more game-like and engaging.

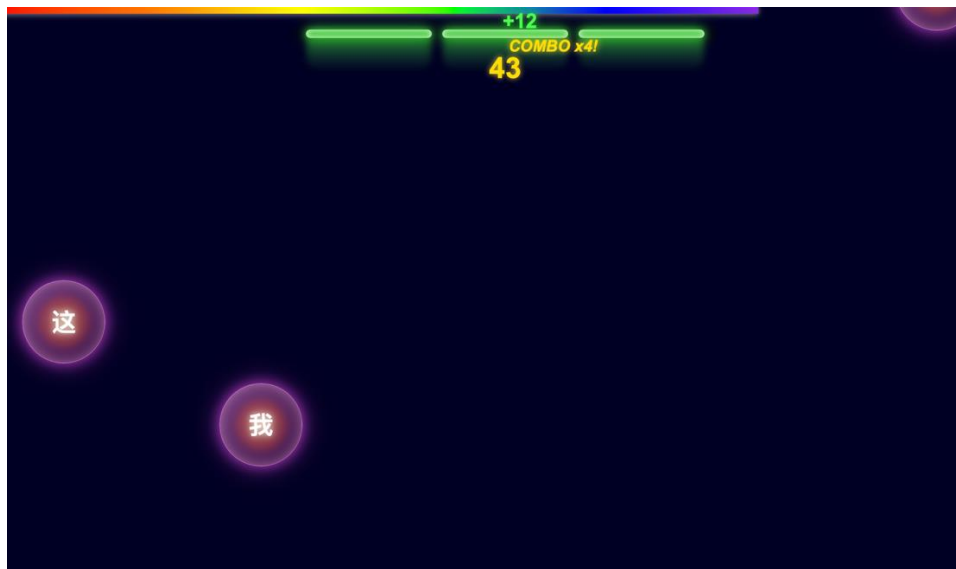


Figure 5. PinyinDash's Time Mode practice screen.

The educational purpose of Time Mode differs from Learning Mode. Time Mode supports rapid retrieval and post-learning review, but it should not be interpreted as complete evidence of tone-marked mastery because learners do not type tone marks during the timed response. Therefore, Learning Mode remains the main space for careful pinyin-tone learning, while Time Mode provides review, motivation, and light pre-exposure to items that may later be practiced more formally.

Brief Dual-Mode Linkage

The two modes are connected at the classroom-practice level rather than treated as isolated mini-games. Students can first use Learning Mode for careful pinyin-tone practice and then use Time Mode for faster review. Time Mode may also help learners become familiar with some items before formal practice, but it does not replace Learning Mode as evidence of pinyin-tone learning. This distinction keeps the prototype's learning flow clear: careful input first, retrieval-oriented review afterward. In this article, the linkage is therefore described only as a learning-flow design, not as a detailed algorithmic mechanism.

Classroom Flow

The prototype also includes progress indicators, feedback, score, wrong-item support, and review cues. These features help students understand their practice status without needing to see the technical scheduling logic. From the learner's perspective, the flow is simple: learn a small group, type pinyin, receive feedback, play a timed review round, and return to weak items when necessary.

This structure makes PinyinDash suitable for short classroom sessions. A teacher can introduce or review target words, students can complete a short Learning Mode activity, and Time Mode can be used as a high-energy retrieval activity. Because the system uses the teacher's own word bank, the game can remain connected to the course syllabus and assessment content. This classroom alignment is central to the present study because the article evaluates PinyinDash as a feasible classroom prototype rather than as a complete standalone learning platform

■ Methods

Research Design

The study used a one-group pre-test/post-test classroom study design. This design was selected because the study aimed to evaluate the classroom feasibility of a newly developed prototype and to examine preliminary evidence of learner motivation and achievement. The design is appropriate for an early-stage prototype evaluation, but it does not allow strong causal claims because it lacks a separate control group. This design aligns with the project's developmental status. PinyinDash was a working classroom prototype, not a mature commercial platform. At this stage, the first empirical question was whether learners could use the game during ordinary class time, whether teacher-selected content could be practiced through the system, and whether the achievement and motivation results were promising enough to justify a larger controlled study. The design emphasizes feasibility and preliminary evidence rather than definitive effectiveness.

Participants and Context

Participants were 30 grade 10 learners at a private high school in Chiang Mai, Thailand. The learners were enrolled in a Chinese course using content aligned with the HSK Level 2 learning level. The intervention was implemented during regular Chinese classes rather than in an artificial laboratory setting. This classroom context was important because the study aimed to examine whether the prototype could be integrated into ordinary instruction.

The instructional sequence included teacher instruction on the target content, a pre-test, 2 weeks of PinyinDash practice, and a post-test. The two-week intervention included eight regular Chinese class sessions. In each session, learners used PinyinDash for approximately 10-15 minutes. The usual classroom sequence was to use Learning Mode first to learn and review the target character-pinyin items, followed by

Time Mode for consolidation through timed retrieval practice. During the two-week intervention, the teacher did not continue reteaching or reviewing the same target item pool; ordinary instruction moved on to later course content, while PinyinDash provided practice on the previously taught items. The teacher-taught item pool served as the basis for both the game and assessment content. All 30 learners completed the pre-test, participated in the full intervention sequence, and completed the post-test, and all results are reported in aggregate form.

The choice of a high school classroom context is important for interpreting the study. Thai learners of Chinese often study characters, pinyin, pronunciation, and meaning within limited weekly class time. Teachers need practice activities that are short, repeatable, and easy to connect with current lessons. The intervention was therefore not designed as a stand-alone self-study course. It was designed as a classroom supplement that could be inserted after the teacher's explanation and before a later assessment.

Intervention Procedure

The classroom intervention followed four stages. First, before the game-practice period, the teacher introduced the target Chinese characters and pinyin forms through normal classroom instruction. Second, learners completed the pre-test to measure their initial performance on the target learning content. Third, learners used PinyinDash across eight class sessions during a two-week practice period. Practice normally began with Learning Mode for slower pinyin-tone learning and review, and then moved to Time Mode for consolidation, rapid retrieval, and game-based practice. Fourth, learners completed the post-test and the ARCS-based motivation questionnaire.

The intervention was designed as an extension of classroom instruction rather than as a replacement for the teacher. This distinction matters because digital tools are often most useful when they provide additional practice, feedback, and motivation around teacher-selected content. In the present study, PinyinDash provided a structured practice environment for content that had already been introduced in class.

During gameplay, learners practiced the same general item pool in both modes. Learning Mode was used for slower pinyin-tone input and correction, while Time Mode provided post-learning review, rapid memory retrieval under game pressure, and light pre-exposure to selected items. Although the teacher retained the flexibility to respond to the classroom rhythm, the typical implementation placed Learning Mode before Time Mode, so that learners first carefully reviewed items and then consolidated them in the timed game environment. This combination of a short practice window and a consistent mode sequence is one reason the prototype was suitable for this classroom implementation.

Research Instruments

Learners' motivation, which was measured using a Thai-language ARCS-based questionnaire after the intervention. The questionnaire contained 16 closed-ended items organized into four dimensions:

Attention (items 1-4), Relevance (items 5-8), Confidence (items 9-12), and Satisfaction (items 13-16). Each dimension contained four items. Responses were given on a five-point Likert scale ranging from 1 = strongly disagree to 5 = strongly agree. For descriptive reporting, responses within each ARCS dimension were summarized using frequencies and percentages. Dimension means were retained only as supplementary descriptive indicators, with higher scores indicating stronger perceived motivation. Internal consistency was evaluated using Cronbach's alpha for the 16 ARCS items overall. The questionnaire was administered after learners had experienced the prototype rather than before use, because the purpose was to measure their perceived motivation toward the actual learning activity. Attention items were interpreted in relation to the game interface and interaction style. Relevance items were interpreted in relation to the connection between game content and classroom learning. Confidence items reflected whether learners felt able to succeed during practice. Satisfaction items reflected whether the game produced a sense of progress, reward, or worthwhile completion.

Another tool is the learners' learning achievement, which was evaluated using a pre-test and a post-test. Each test had a maximum score of 20 points and contained 20 items. Ten items were multiple-choice questions that included pinyin with tone information in the answer options, and ten items were pinyin-writing items requiring learners to write the pinyin corresponding to given Chinese characters. Each item was worth one point. For the pinyin-writing items, learners wrote standard pinyin with tone marks for the given Chinese characters, such as *mā*, rather than tone numbers. Each response was judged as correct or incorrect according to a researcher-prepared answer key; no partial credit was given because the pinyin form was treated as a complete answer. The school's Chinese teacher first checked the responses, and the researcher then reviewed the judgments. For polyphonic characters, any correct pronunciation was accepted as correct. The two tests were not identical repeated forms. Instead, they were designed according to the same assessment blueprint, used the same item-type balance, were intended to be similar in difficulty, and sampled Chinese characters from the same teacher-taught item pool. The two-part test format was chosen to reflect both recognition and production. Multiple-choice items can show whether learners recognize the correct tone-marked pinyin among alternatives, while pinyin-writing items require learners to produce the complete pinyin form from a Chinese character prompt. To reduce the possibility that learners would improve simply by remembering the exact pre-test items, words appearing in prompts and options were replaced where appropriate in the post-test. These replacement words were selected from the same classroom item pool, so the two forms remained aligned with the same target content while avoiding direct item repetition. These procedures provided content validity evidence because the test content, classroom instruction, game word bank, and target construct were aligned around Chinese character-pinyin mapping. The two components were combined into a total achievement score because both measured complementary aspects of the same target construct: connecting Chinese characters with corresponding pinyin forms. The total score was therefore interpreted as an overall classroom-aligned indicator of character-pinyin learning.

Internal consistency was examined for the dichotomously scored achievement tests. The pre-test showed acceptable internal consistency, Cronbach's $\alpha = .740$, and the post-test also showed acceptable internal consistency, Cronbach's $\alpha = .733$.

Data Analysis

Descriptive statistics were calculated for pre-test and post-test total scores. A Shapiro-Wilk test was used to examine the normality of difference scores. A paired-samples t-test was used to compare pre-test and post-test achievement. The mean gain and paired-samples effect size d_z were reported to describe the magnitude of change. For the paired-samples comparison, d_z was calculated as $t/\sqrt{N}$, and its absolute value was interpreted as the magnitude of the effect. Component-level descriptive statistics were also calculated for the multiple-choice recognition items and the pinyin-writing production items. Internal consistency for the dichotomously scored achievement tests was examined using Cronbach's α . For motivation, response frequencies and percentages were calculated for each ARCS dimension, and dimension means were retained as supplementary descriptive indicators. Cronbach's α was used to evaluate overall internal consistency.

The statistical analysis was kept aligned with the one-group classroom study design. The paired-samples t-test was suitable because the same learners completed the pre-test and post-test. The paired-samples effect size d_z was reported to provide an effect-size estimate, but it is interpreted cautiously because there was no comparison group. The motivation analysis focused on descriptive patterns and reliability rather than prediction. This choice reflects the article's purpose: to report whether the prototype is promising in classroom use and to identify design implications for future controlled studies.

Detailed in-game log data were not collected during this classroom implementation. Therefore, the present analysis focuses on achievement outcomes and post-use motivation responses rather than behavioral process data such as mode-use frequency, item-level attempts, response times, or wrong-item review actions.

■ Research Results

Learning Motivation Results

Table 2 presents the response frequency and percentage distribution for the ARCS motivation questionnaire. The overall 16-item ARCS scale showed high internal consistency, Cronbach's $\alpha = .928$. Across the four dimensions, most responses were either agree or strongly agree. The combined agree/strongly agree percentage was highest for Relevance (95.0%), followed by Satisfaction (92.5%), Attention (88.3%), and Confidence (85.8%). These results suggest that learners generally perceived

PinyinDash as relevant to their classroom learning, satisfying to use, attention-supporting, and confidence-supporting.

As supplementary descriptive indicators, the dimension means were also high: Satisfaction ($M = 4.66$), Relevance ($M = 4.59$), Attention ($M = 4.51$), and Confidence ($M = 4.44$).

The pattern across the four dimensions provides useful design information. The high proportion of agree/strongly agree responses for Satisfaction suggests that learners experienced the activity as rewarding or worthwhile. The high proportion of agree/strongly agree responses for Relevance suggests that learners did not perceive the game as unrelated entertainment; instead, they recognized a connection between PinyinDash and their Chinese learning. Attention was also high, which is consistent with the use of animated falling items, feedback, score, and short rounds. Confidence showed the lowest agree/strongly agree percentage among the four dimensions, although it remained high, which may reflect the difficulty of producing pinyin accurately, especially when tone information is involved.

Taken together, the motivation results support the feasibility of continuing to use and refine the prototype in classroom settings. They do not prove that motivation caused achievement gains, but they show that learners' affective response was positive after the intervention. This matters because character-pinyin practice requires repeated effort. A tool that improves practice opportunities but is experienced as frustrating or irrelevant would be unlikely to survive in regular classroom use.

Table 2.

ARCS response distribution by dimension

ARCS dimension	N	1 n	2 n	3 n	4 n	5 n	4-5 n
	Response	(%)	(%)	(%)	(%)	(%)	(%)
Attention	120	0 (0.0%)	0 (0.0%)	14 (11.7%)	31 (25.8%)	75 (62.5%)	106 (88.3%)
Relevance	120	0 (0.0%)	1 (0.8%)	5 (4.2%)	36 (30.0%)	78 (65.0%)	114 (95.0%)
Confidence	120	0 (0.0%)	3 (2.5%)	14 (11.7%)	30 (25.0%)	73 (60.8%)	103 (85.8%)
Satisfaction	120	0 (0.0%)	0 (0.0%)	9 (7.5%)	23 (19.2%)	88 (73.3%)	111 (92.5%)

Note. Each ARCS dimension contained four items; therefore, each dimension included 120 item responses from 30 learners. Scale: 1 = strongly disagree, 5 = strongly agree. The 4-5 column combines agree and strongly agree responses.

Learning Achievement Results

Table 3 presents the component-level and total descriptive results for the achievement test scores. The Shapiro-Wilk test indicated that the difference scores were approximately normally distributed, $W=.981$,

$p = .841$. A paired-samples t-test showed a statistically significant increase from pre-test to post-test, $t(29) = -10.85$, $p < .001$. The negative sign reflects the subtraction order used in the paired comparison. The mean gain was 5.37 points on the 20-point scale, and the effect size was large, $|d_z| = 1.98$.

Because the two tests followed the same assessment blueprint but did not repeat the same items, this comparison is interpreted as change across two classroom-aligned forms rather than as performance on an identical repeated test.

The descriptive pattern is also educationally meaningful. The total mean score increased from 65.00% to 91.85% of the maximum score. The median score increased from 13.00 to 19.00, suggesting that the improvement was not limited to a small number of learners. The standard deviation decreased from 3.19 to 2.19, which indicates that post-test performance was not only higher on average but also more concentrated near the upper end of the scale. In this classroom study, this pattern suggests broad improvement during the PinyinDash practice period on a target item pool introduced before the pre-test.

The result should still be read within the scope of the assessment. The tests measured the teacher-taught character-pinyin item pool rather than general Chinese proficiency. Therefore, the gain indicates improvement across two classroom-aligned test forms covering the target learning content used in the course and game intervention. This is appropriate for prototype evaluation because the purpose of PinyinDash is to support classroom-aligned practice, but it should not be generalized to all aspects of Chinese vocabulary or communicative ability.

Component-level descriptive results clarified the mixed assessment structure. The multiple-choice recognition component increased from $M = 8.67$ ($SD = 1.30$) to $M = 9.70$ ($SD = 0.65$), with a mean gain of 1.03 points out of 10. The pinyin-writing production component increased from $M = 4.33$ ($SD = 2.41$) to $M = 8.67$ ($SD = 1.97$), with a mean gain of 4.33 points out of 10. Therefore, improvement appeared in both test formats, but the larger gain was found in productive pinyin writing.

The achievement test forms showed acceptable internal consistency, with Cronbach's $\alpha = .740$ for the pre-test and .733 for the post-test.

Table 3.

Component-level descriptive results for achievement scores

Component	N	Max score	Pre-test M	Pre-test SD	Post-test M	Post-test SD	Mean Gain
Multiple-choice	30	10	8.67	1.30	9.70	0.65	1.03
Pinyin-writing	30	10	4.33	2.41	8.67	1.97	4.33
Total achievement	30	20	13.00	3.19	18.37	2.19	5.37

■ Discussion

Prototype Feasibility

The results provide preliminary classroom evidence for the feasibility of PinyinDash as a dual-mode digital game prototype. The learning gain from pre-test to post-test suggests that learners improved on the teacher-taught character-pinyin content during the intervention period. The motivation results suggest that learners responded positively to the game experience. Taken together, these findings support continuing to develop and evaluate the prototype. The prototype structure provides a plausible design-based explanation for why the tool may be useful in classroom practice, although the present study did not collect log data to test the separate contribution of each feature. Learning Mode creates a low-pressure space for pinyin-tone input and correction. The five-item group structure keeps practice manageable, while the retry queue prevents incorrect items from being ignored. This aligns with retrieval practice because learners must actively produce pinyin forms rather than simply identify them. It also supports confidence because the learner can complete a small group through repeated attempts rather than facing a long list of disconnected items.

The five-item group structure is especially important for classroom feasibility. A long list of characters can make beginners feel that the task is endless, while a very small task may not provide enough repetition. Five items create a short cycle that can be completed, reviewed, and repeated within a brief class period. This supports the teacher's need for manageable classroom activities and the learner's need for visible completion. It also makes progress feedback easier to understand because learners can see movement from a small current set toward broader mastery. Time Mode contributes a different kind of value. The falling-meteor mechanic, timer, lives, score, and combo feedback create urgency and engagement. These mechanics can make repeated pinyin retrieval feel less like mechanical drilling and more like a challenge. The mode is especially useful after initial instruction because it encourages learners to retrieve pinyin from memory under bounded time pressure. However, the design should still be interpreted carefully: Time Mode uses toneless pinyin, so it supports rapid form retrieval rather than complete tone-marked mastery. The prototype is strongest when Learning Mode and Time Mode are used together.

This dual structure also helps manage the tension between accuracy and time-pressured retrieval. If tone-marked pinyin input were forced into every timed response, some learners might spend more effort fighting the input method than retrieving the target form. If tones were removed from all practice, the game would avoid a central difficulty of pinyin learning. PinyinDash resolves this tension by assigning different learning functions to different modes. Accuracy is protected in Learning Mode, while review, repeated retrieval, and light pre-exposure are emphasized in Time Mode. The design avoids treating one mode as sufficient for all learning purposes. The same boundary is important for interpreting pre-exposure and adaptive difficulty. Pre-exposure should be understood as a way to order future Learning Mode practice, not as implicit evidence that an item has been learned. Similarly, the 85 percent target belongs to Time Mode meteor-level pacing and challenge control; it was not used as a mastery threshold and was not separately validated in this classroom study.

Motivational Interpretation

The ARCS results are consistent with the design rationale summarized in Table 1. The high agree/strongly agree percentage for Attention suggests that learners responded positively to the falling-meteor review environment, dual time pressure, and immediate visual feedback. These features were designed to make repeated pinyin retrieval more dynamic while keeping the core game action connected to pinyin responses.

The high agree/strongly agree percentage for Relevance is also important because PinyinDash was designed around teacher-aligned word banks and a shared item pool across Learning Mode and Time Mode. This structure helped connect slow learning, fast review, and classroom-taught vocabulary. Light Time Mode pre-exposure may also have made some future Learning Mode items more familiar, although it should not be interpreted as evidence that those items had already been mastered.

Confidence showed the lowest agree/strongly agree percentage among the four dimensions, although it remained high. This pattern is understandable because productive pinyin retrieval is demanding for beginning learners. The prototype addressed this challenge through five-item Learning Mode groups, mistake cards, retry opportunities, and visible mastery states. These features make the task more manageable and help learners treat errors as part of guided practice rather than as final failure.

Satisfaction showed the highest agree/strongly agree percentage among the four dimensions. This may reflect the way successful responses are made visible through success cards, meteor destruction, score and combo rewards, group-completion review, visible progress indicators, and the optional local leaderboard. These rewards are tied to successful pinyin practice, so the game reward and the learning action remain connected. However, the ARCS data should still be interpreted as learners' overall response to the prototype, not as evidence that any single feature independently caused the ratings.

Achievement Interpretation

The achievement gain is encouraging but should be interpreted as preliminary. The learners received teacher instruction on the target item pool before the PinyinDash practice period, and the study did not include a control group. During the two-week intervention, regular teaching continued with later course content rather than additional reteaching of the same target items. Therefore, the result still cannot prove that PinyinDash alone caused the improvement, but it is more accurate to say that learners improved after a classroom sequence that included prior teacher instruction and eight short PinyinDash practice sessions. Even with this caution, the result is meaningful for prototype evaluation. A digital learning tool must first demonstrate that it can be integrated into classroom practice and that learners can use it in a way that supports positive learning change. The large mean gain and high motivation ratings suggest that PinyinDash is promising enough for more rigorous future research using stronger classroom designs and richer process data.

The assessment format also matters for interpreting the gain. Because half of each test required pinyin-writing responses with tone information and did not allow partial credit, the post-test improvement cannot be explained only as better recognition of answer choices. It suggests improvement in the type of active pinyin-tone retrieval that the prototype is designed to support. The two forms were designed with the same blueprint and target item pool, but the prompt and option words were replaced where appropriate to reduce memorization of exact test items. The component-level descriptive results now clarify that improvement appeared in both recognition and productive pinyin-writing components, with a larger gain in productive pinyin writing. Future controlled studies should examine whether game practice contributes not only to recognition but also to productive pinyin spelling with tone information.

Implications for Chinese Character-Pinyin Learning

The findings also suggest a broader design implication for Chinese learning tools. Character-pinyin practice should not be limited to recognition. Beginning learners need opportunities to produce pinyin forms, notice errors, and repeat weak items. At the same time, repeated production needs to be made motivationally sustainable. PinyinDash addresses this balance by combining careful input practice with a review-oriented game mode. This dual-mode structure may be useful for other language-form learning contexts where learners need both accuracy and repeated memory retrieval.

For teachers, the main implication is that a digital game can serve as a targeted practice layer rather than a separate entertainment activity. The teacher remains responsible for introducing characters, explaining pronunciation, and situating words in meaningful use. The game then provides additional opportunities for learners to retrieve pinyin forms and receive immediate feedback. This division of roles is realistic for school contexts because it does not require the technology to carry the entire burden of instruction. It also allows the teacher to adapt the word bank to lessons, tests, or review needs.

For researchers, the study shows why prototype papers should report both design structure and learner outcomes. If only achievement is reported, it is difficult to understand which practice produced the result. If only the prototype is described, it is difficult to judge whether the tool is promising in real classroom use. By combining structure, achievement, and ARCS motivation, the present article provides a fuller early-stage evaluation. The evidence remains preliminary, but it identifies a clear design direction for more rigorous future research.

Comparison with Previous Studies

The findings of the present study are broadly consistent with previous research on digital game-based language learning. Prior studies have suggested that language games can support engagement, vocabulary practice, and learning outcomes when game actions are closely connected to the target learning task (Butler, 2015; Chen & Hsu, 2020; Zhonggen, 2018). The improvement observed in this study is compatible with this view because PinyinDash required learners to respond directly to Chinese character-

pinyin prompts rather than interact with game elements unrelated to the learning content. The results also align with vocabulary game studies reporting positive or supportive patterns for vocabulary achievement, motivation, and word knowledge (Li, 2021; Mokhtar & Rahelan, 2024; Rasti-Behbahani & Shahbazi, 2020).

The achievement pattern is also consistent with retrieval-practice and spaced-review research. Previous studies have shown that active retrieval can strengthen later retention more effectively than passive exposure or restudy (Roediger & Karpicke, 2006; Karpicke & Roediger, 2008; Karpicke & Blunt, 2011), and that distributed practice can support long-term verbal learning (Cepeda et al., 2006, 2008). PinyinDash reflects these principles by asking learners to produce pinyin forms in Learning Mode and by creating repeated review opportunities through Time Mode. However, the present study should be interpreted as classroom feasibility evidence rather than direct proof of these cognitive mechanisms, because it did not include a control group or item-level process analysis.

The motivation results also correspond to the ARCS-based design rationale. Keller's ARCS model emphasizes Attention, Relevance, Confidence, and Satisfaction as important motivational dimensions in instructional design (Keller, 2010). In this study, all four dimensions were rated highly, suggesting that learners perceived the prototype as engaging, relevant to class, manageable, and satisfying. This is consistent with previous Chinese vocabulary game studies that highlight the motivational potential of gamified vocabulary practice (Rawendy et al., 2017; Ying et al., 2021; Phupunna & Hongsamsibkao, 2024; Zhang & Abdul Salam, 2025). At the same time, Confidence was the lowest of the four ARCS dimensions, which suggests a useful difference from purely motivational interpretations: learners may enjoy the game and find it relevant, but pinyin-tone production remains a demanding learning task that still requires scaffolding.

■ Conclusion

This article presented PinyinDash, a dual-mode digital game prototype for Chinese character-pinyin learning, and reported preliminary classroom evidence from a participating group of Thai high school learners. The prototype combines Learning Mode for accurate pinyin-tone input with Time Mode for post-learning review, rapid memory retrieval, and light pre-exposure through falling-meteor gameplay. It also supports teacher-aligned word banks, immediate feedback, visible progress, wrong-item support, and review support. After two weeks of classroom use, learners showed a statistically significant improvement from pre-test to post-test and reported high motivation across all four ARCS dimensions. These results suggest that PinyinDash is a feasible and motivationally acceptable prototype for supporting Chinese character-pinyin practice. The findings should be treated as preliminary rather than causal, but they provide a strong basis for future controlled studies and continued prototype refinement.

■ Limitations and Recommendations

This study has several limitations. First, the one-group pre-test and post-test design limits causal interpretation. Without a control group, the observed improvement cannot be attributed only to PinyinDash. Second, the sample size was appropriate for an early-stage classroom study but remains limited for broader generalization. Third, although the revision reports component-level descriptive results for multiple-choice recognition and pinyin-writing production, the present study does not conduct item-level difficulty or discrimination analyses. It also does not model the independent contribution of Learning Mode and Time Mode to each test component.

Another limitation is the lack of detailed in-game log data from the classroom implementation. The available results show that learners improved and reported positive motivation after using the prototype, but they do not show which specific game behaviors were most closely related to improvement. Therefore, claims about progress indicators, retry opportunities, wrong-item support, review cues, score feedback, combo rewards, and leaderboard use should be interpreted as design-based explanations rather than feature-level causal evidence. The ARCS results should also be interpreted as learners' overall post-use perceptions rather than evidence that any single feature independently caused motivation.

Future research should use controlled or quasi-experimental classroom designs with larger and more diverse samples. It should also collect detailed usage logs, including Learning Mode attempts, Time Mode accuracy, response time, mode-use duration, wrong-item records, review actions, score and combo records, leaderboard use, and item-level progress changes. Such data would make it possible to examine which parts of the prototype are most closely associated with learning gains and motivation. Future studies could also include delayed post-testing to examine whether character-pinyin learning is retained beyond the immediate classroom intervention.

References

- Bustrillo, H. T., Sumicad, R., & Cuevas, G. C. (2024). Effectiveness of word games in teaching students vocabulary. *Journal of World Englishes and Educational Practices*, 6(3), 7-16. <https://doi.org/10.32996/jweep.2024.6.3.2>
- Butler, Y. G. (2015). The use of computer games as foreign language learning tasks for digital natives. *System*, 54, 91-102. <https://doi.org/10.1016/j.system.2014.10.010>
- Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, 132(3), 354-380. <https://doi.org/10.1037/0033-2909.132.3.354>
- Cepeda, N. J., Vul, E., Rohrer, D., Wixted, J. T., & Pashler, H. (2008). Spacing effects in learning: A temporal ridge line of optimal retention. *Psychological Science*, 19(11), 1095-1102. <https://doi.org/10.1111/j.1467-9280.2008.02209.x>

- Chen, H. J. H., & Hsu, H. L. (2020). The impact of a serious game on vocabulary and content learning. *Computer Assisted Language Learning*, 33(7), 811-832. <https://doi.org/10.1080/09588221.2019.1593197>
- Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science*, 319(5865), 966-968. <https://doi.org/10.1126/science.1152408>
- Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*, 331(6018), 772-775. <https://doi.org/10.1126/science.1199327>
- Keller, J. M. (2010). *Motivational design for learning and performance: The ARCS model approach*. Springer. <https://doi.org/10.1007/978-1-4419-1250-3>
- Li, R. (2021). Does game-based vocabulary learning APP influence Chinese EFL learners' vocabulary achievement, motivation, and self-confidence? *SAGE Open*, 11(1), 21582440211003092. <https://doi.org/10.1177/21582440211003092>
- Mohd Mokhtar, M. A., & Rahelan, A. N. A. (2024). The use of gaming tools on students to improve vocabulary mastery in the classroom. *International Journal of Academic Research in Business and Social Sciences*, 14(9), 1223-1233. <https://doi.org/10.6007/IJARBS/v14-i9/22741>
- Phupunna, S., & Hongsamsibkao, N. (2024). Enhancing Chinese vocabulary memorization skills through Blooket games combined with active learning for first-year students in the Chinese language teaching program at Rajabhat Mahasarakham University. *Journal of Education and Learning*, 13(5), 191. <https://doi.org/10.5539/jel.v13n5p191>
- Rasti-Behbahani, A., & Shahbazi, M. (2020). Investigating the effectiveness of a digital game-based task on the acquisition of word knowledge. *Computer Assisted Language Learning*, 35(8), 1920-1945. <https://doi.org/10.1080/09588221.2020.1846567>
- Rawendy, D., Ying, Y., Arifin, Y., & Rosalin, K. (2017). Design and development game Chinese language learning with gamification and using mnemonic method. *Procedia Computer Science*, 116, 61-67. <https://doi.org/10.1016/j.procs.2017.10.009>
- Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, 17(3), 249-255. <https://doi.org/10.1111/j.1467-9280.2006.01693.x>
- Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, 15(1), 20-27. <https://doi.org/10.1016/j.tics.2010.09.003>
- Ying, Y., Susilo, P. M., Mei, F. R., & Rahardjanti, T. (2021). The role of the Mandamonic games in supporting Mandarin learning at elementary school. *Journal of Physics: Conference Series*, 1764(1), 012107. <https://doi.org/10.1088/1742-6596/1764/1/012107>
- Zhonggen, Y. (2018). Differences in serious game-aided and traditional English vocabulary acquisition. *Computers & Education*, 127, 214-232. <https://doi.org/10.1016/j.compedu.2018.07.014>
- Zhang, J., & Abdul Salam, S. N. (2025). Research on the design of Chinese vocabulary educational games based on the ARCS model. *Multidisciplinary Science Journal*, 7(10), 2025385. <https://doi.org/10.31893/multiscience.2025385>