

Immediate feedback in gamified online Chinese vocabulary quizzing: Effects on perceived cognitive load and strategic vocabulary learning

 Yan Song¹,  Sirirat Petsangsri^{2*},  Thanin Rattana-Olan³,  Jianming Cui⁴

¹School of Industrial Education and Technology, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand; liujiapiao74robin@126.com (Y.S.).

^{2,3}School of Industrial Education and Technology, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand; sirirat.pe@kmitl.ac.th (S.P.) thanin.ra@kmitl.ac.th (T.R.O.).

⁴School of Humanities and International Studies, Zhejiang University of Water Resources and Electric Power, Zhejiang, China; cui_jm@zjweu.edu.cn (J.C.).

Abstract: Online Chinese-as-a-second-language (CSL) learners often experience substantial cognitive demands during vocabulary study because they must rapidly encode and retrieve unfamiliar form–meaning mappings under constrained working-memory resources. Gamified quizzing platforms (e.g., Quizizz) commonly provide immediate feedback (verification plus brief explanatory feedback), yet evidence remains limited regarding how this feature shapes learners' perceived cognitive load and vocabulary learning strategy use in online CSL contexts. This randomized study investigated whether embedding immediate feedback into weekly online Chinese vocabulary quizzes over a six-week HSK preparation course reduces learners' perceived task workload and increases their strategic vocabulary learning. Sixty undergraduates were sampled using gender-stratified procedures and allocated by block randomization to Immediate Feedback, Leaderboards, or non-gamified Control conditions (n=20 per condition). Following each weekly quiz, participants completed the NASA Task Load Index (NASA-TLX) and an adapted Vocabulary Learning Strategies Questionnaire. Participant-level composite scores were analyzed using path-analytic regression with dummy-coded conditions (control as the reference). Immediate feedback significantly reduced perceived cognitive load ($\beta=-0.42$, SE = 0.08, $t = -5.25$, $p < .001$) and significantly increased reported strategy use ($\beta=0.55$, SE=0.07, $t=7.86$, $p < .001$). Interview themes aligned with the quantitative results, emphasizing uncertainty reduction, lower anxiety and frustration, increased perceived control, and more deliberate monitoring and planning when explanatory feedback was provided. These findings indicate that immediate feedback in gamified assessment can serve as a cognitively significant scaffold supporting strategic vocabulary learning during online CSL retrieval practice.

Keywords: Cognitive load, Gamified Quizzes, Immediate feedback, Online Chinese learning, Vocabulary learning strategies.

1. Introduction

Online Chinese-as-a-second-language (CSL) provision has expanded rapidly, yet sustaining effective learning processes in online settings remains challenging. Compared with face-to-face contexts, online CSL learners often encounter dense information with limited instructor immediacy and heightened demands for self-regulation. Vocabulary learning is a particularly intensive component of CSL development because it requires repeated retrieval and consolidation of unfamiliar form–meaning mappings, often involving orthographic and phonological novelty. When intrinsic task demands are high, instructional and assessment designs that reduce unnecessary cognitive burden while supporting effective learning behaviors become especially important.

Gamified quiz platforms (e.g., Quizizz) have been widely adopted to support engagement, repeated retrieval practice, and formative assessment in vocabulary learning. However, gamification is not a unitary intervention. Platforms typically bundle multiple design elements, feedback, points, timers, leaderboards, and other mechanics that may operate through different cognitive, motivational, and self-regulatory pathways. Treating gamification as an undifferentiated package can obscure the contribution of specific features that may be educationally beneficial (or detrimental) for particular language-learning tasks, limiting both theoretical interpretation and practical design guidance.

One ubiquitous and instructionally meaningful element is immediate feedback, typically delivered via (a) an immediate correctness indicator and (b) a brief explanation or corrective prompt presented immediately after each response. From a cognitive load perspective, immediate explanatory feedback may reduce extraneous processing by minimizing uncertainty and improving the efficiency of error diagnosis during timed retrieval practice. From a self-regulated learning perspective, timely correctness information can serve as a monitoring cue that helps learners identify gaps, calibrate performance, and adjust subsequent study strategies.

Despite the prevalence of immediate feedback in gamified quizzing, empirical evidence remains limited regarding its cognitive and strategic consequences in online CSL vocabulary quizzing. Prior work has often emphasized engagement or broad “gamification effects,” rather than isolating specific platform features and examining their influence on learners' perceived cognitive load and vocabulary learning strategy use during retrieval-based vocabulary tasks. Addressing this gap, the present study examines whether embedding immediate feedback (verification plus brief explanatory feedback) into weekly gamified Chinese vocabulary quizzes within a six-week online HSK preparation course is associated with (i) lower perceived cognitive load and (ii) greater reported use of vocabulary learning strategies. Accordingly, we test two hypotheses: (H₁) immediate feedback is negatively associated with perceived cognitive load, and (H₂) immediate feedback is positively associated with vocabulary learning strategy use.

1.1. Research Questions

RQ1: To what extent does immediate feedback relate to learners' perceived cognitive load in online Chinese vocabulary quizzing?

RQ2: To what extent does immediate feedback relate to learners' reported use of vocabulary learning strategies during online Chinese vocabulary quizzing?

1.2. Research Objectives

The present study investigates the role of immediate feedback, defined as item-level verification with brief explanatory information, embedded within gamified online Chinese vocabulary quizzes. Specifically, it examines whether providing such feedback during weekly, retrieval-based quizzes is associated with lower perceived cognitive load and greater use of vocabulary learning strategies among online Chinese-as-a-second-language learners.

Accordingly, the study pursues two objectives:

1. To determine the extent to which immediate feedback is associated with learners' perceived cognitive load during online Chinese vocabulary quizzes.
2. To determine the extent to which immediate feedback is associated with learners' reported use of vocabulary learning strategies during online Chinese vocabulary quizzing.

1.3 Hypotheses

H₁: Immediate feedback is associated with lower perceived cognitive load.

H₂: Immediate feedback is associated with higher reported vocabulary learning strategy use.

2. Literature Review and Hypothesis Development

2.1. Gamified Quizzing as Formative Assessment for Vocabulary Learning

Gamification is commonly conceptualized as the integration of game design elements (e.g., points, feedback, leaderboards, time constraints) into non-game contexts to influence engagement and persistence [1-3]. In second-language learning contexts, quiz-based gamified platforms can be particularly compatible with vocabulary acquisition because they support frequent, low-stakes retrieval practice, a mechanism widely linked to durable learning through test-enhanced learning effects [4, 5].

However, increased activity or engagement does not necessarily translate into improved learning processes. For CSL vocabulary learning, learners must encode and retrieve unfamiliar form–meaning mappings under working-memory constraints, which can elevate cognitive demands. Consequently, the pedagogical value of gamified quizzing depends not only on motivational affordances but also on whether specific design features promote cognitive efficiency and strategic self-regulation during retrieval-based practice. From this perspective, gamification should be treated as a constellation of distinct features rather than a unitary intervention; isolating feature-level effects is necessary to generate interpretable evidence and actionable design guidance [6].

2.2. Immediate Feedback and Learning: Timing and Informativeness

A central feature of many digital quiz platforms is immediate feedback, often delivered as correctness verification (correct/incorrect) and, in more instructionally rich implementations, brief explanatory or corrective information. Formative feedback is most effective when it helps learners recognize discrepancies between current performance and target standards and provides information that can guide subsequent improvement [7, 8]. Meta-analytic and theoretical work further suggests that feedback can influence learning outcomes depending on its characteristics, including timing and informational content [9, 10].

In vocabulary retrieval tasks, the timing of feedback is consequential because corrective information provided while the learner's task representation remains active may facilitate more efficient updating of form–meaning mappings and reduce the likelihood that errors persist or are reinforced through repeated retrieval attempts. Moreover, feedback informativeness matters: while verification feedback supports performance calibration, brief explanations can additionally facilitate error diagnosis by clarifying why a response is incorrect and how to repair the mapping. Given the flow demands of gamified quizzing, explanatory feedback is typically concise; nevertheless, even brief explanatory cues may meaningfully shape cognitive processing during fast-paced online vocabulary practice.

In the present study, immediate feedback is operationalized as item-level correctness information plus a brief explanation presented immediately after each response, without allowing retries within the same quiz session.

2.3. Cognitive Load Theory and Why Immediate Feedback May Reduce Perceived Load

Cognitive Load Theory (CLT) posits that learning is constrained by the limited capacity of working memory; instructional design should therefore manage total cognitive load by minimizing avoidable extraneous demands and supporting processing that contributes to schema acquisition [11-13]. CSL vocabulary learning can impose a high intrinsic load because learners must process novel orthographic forms, pronunciations, and semantic distinctions simultaneously. Under such conditions, design features that reduce unnecessary processing demands may be particularly influential.

Immediate feedback is expected to reduce perceived task workload by limiting extraneous processing through at least three complementary mechanisms.

1. Uncertainty reduction: Immediate confirmation minimizes the need to hold tentative answers in mind or wait until quiz completion to resolve correctness, thereby reducing ongoing mental burden.
2. Error-diagnosis efficiency: Brief explanations streamline error interpretation and reduce the need for post-hoc searching or reconstructing the cause of mistakes.

3. Affective load mitigation: Rapid clarification of errors may lower frustration and anxiety, affective states that can consume cognitive resources and inflate perceived workload.

Consistent with CLT-based arguments and prior work on cognitive load in gamified contexts [14], immediate feedback should therefore be associated with lower perceived cognitive load during online CSL vocabulary quizzing.

Accordingly, immediate feedback should be negatively associated with perceived task load (H_1).

2.4. Feedback, Self-Regulation, and Vocabulary Learning Strategies

In second-language acquisition research, vocabulary learning success is strongly linked to learners' deployment of strategies across multiple domains, including memory, cognitive, metacognitive, and social/affective strategies [15-18]. Strategies such as rehearsal, elaboration, imagery, systematic review planning, progress monitoring, and targeted attention to weak items reflect learners' capacity to regulate learning conditions and manage the demands of vocabulary acquisition.

Feedback is also a core mechanism in self-regulated learning (SRL) frameworks. Theoretical syntheses describe feedback as information that learners use to monitor performance, evaluate progress, and adapt tactics and effort [19]. Similarly, formative assessment models emphasize that effective feedback provides information that can be translated into action through goal-setting, monitoring, and strategy adjustment [20]. In online quizzing environments, immediate feedback may therefore function as a *salient monitoring signal*, a prompt indicator of what is known versus uncertain, that supports strategic decisions such as allocating additional practice, shifting from shallow repetition to deeper encoding, or maintaining records of recurring errors.

Accordingly, by strengthening performance monitoring and enabling timely adjustment, immediate feedback is expected to be associated with greater reported use of vocabulary learning strategies. Thus, immediate feedback should be positively associated with reported strategy use (H_2).

2.5. Conceptual Model

Conceptual model of hypothesized relationships

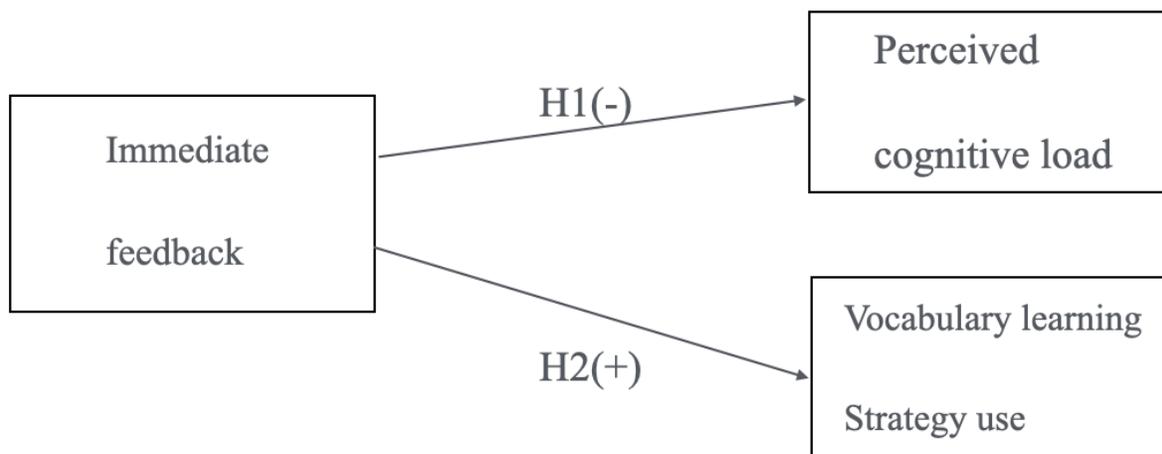


Figure 1.
Conceptual Model.

3. Methodology

3.1. Research Design and Setting

This study employed a randomized, multi-arm experimental design implemented in an online HSK preparation course delivered at a Chinese university over six weeks. Participants completed weekly online Chinese vocabulary quizzes hosted on Quizizz. Although the broader parent design included three conditions, Immediate Feedback, Leaderboards, and a non-gamified Control, the present manuscript foregrounds the immediate feedback feature as the focal design element while retaining the full randomized structure to preserve design integrity and comparability across conditions.

3.2. Participants, Sampling, and Randomization

Participants were 60 undergraduate students enrolled in the online HSK preparation course. A gender-stratified sampling procedure was used to maintain proportional representation, followed by block randomization within strata to allocate participants equally to the three conditions (n=20 per condition). Baseline equivalence was examined using a pre-intervention HSK-based vocabulary test, with comparable mean scores observed across groups (see Table 1 in the manuscript).

Table 1.
Participant characteristics and baseline equivalence across conditions

Characteristic	Immediate Feedback (n = 20)	Leaderboards (n = 20)	Control (n = 20)	Total (n = 60)
Female, n (%)	12 (60%)	12 (60%)	12 (60%)	36 (60%)
Male, n (%)	8 (40%)	8 (40%)	8 (40%)	24 (40%)
Age (years), M (SD)	19.9 (0.9)	20.1 (1.0)	20.0 (1.1)	20.0 (1.0)
Pre-test vocabulary score, M (SD)	33.6 (5.7)	33.4 (5.9)	33.3 (6.0)	33.4 (5.8)

Note: Pre-test refers to the HSK-based vocabulary test administered prior to the intervention.

3.3. Conditions and Operationalization of Immediate Feedback

Eleven participants completed the same weekly quizzes under standardized implementation controls. Specifically, quiz content and delivery were held constant across conditions, including identical item sets, quiz forms and ordering, time limits, device rules, instructor, and course materials. Scoring was accuracy-based only (correct = 1; incorrect = 0), rescaled to a 0–100 score, with no time- or streak-based bonuses.

The conditions were defined as follows:

Immediate Feedback condition: After each response, learners received item-level correctness verification and a brief explanatory/corrective statement presented immediately. Learners could not retry items within the same quiz session.

Control condition (non-gamified): Learners did not receive item-level explanatory feedback during quizzing; instead, they received only overall quiz scores upon quiz completion.

Leaderboards condition: This condition was included as part of the parent randomized design. In the present manuscript, it is retained as a design-control comparator, but the primary theoretical and interpretive focus is on isolating the immediate feedback feature.

3.4. Procedure and Timeline

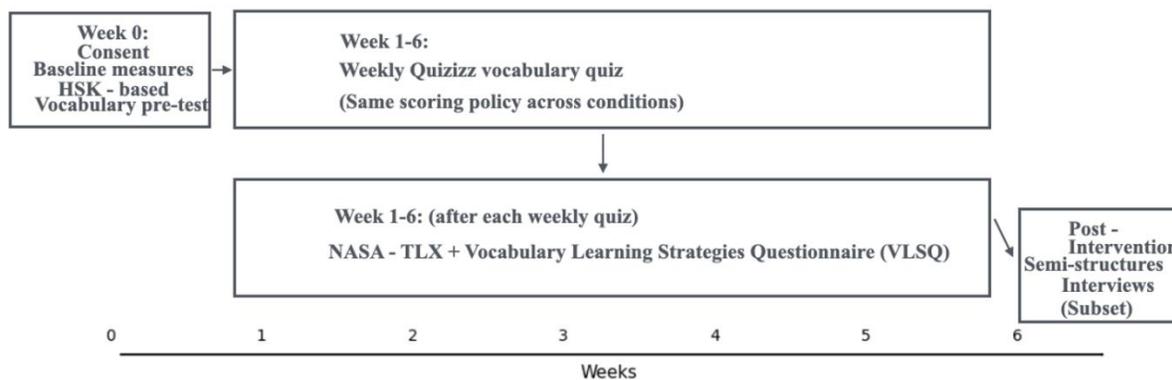


Figure 2.
Study procedure and timeline (six weeks).

The study followed a structured six-week timeline:

1. Week 0 (baseline): Participants provided consent and completed baseline measures, including an HSK-based vocabulary pre-test.
2. Weeks 1–6 (intervention phase): Participants completed one weekly Quizizz vocabulary quiz according to their assigned condition, with consistent scoring policies and implementation controls maintained across groups.
3. Weeks 1–6 (post-quiz measures): Immediately following each weekly quiz, participants completed (i) the NASA Task Load Index (NASA-TLX) and (ii) the adapted Vocabulary Learning Strategies Questionnaire, as specified in the study protocol.
4. Post-intervention qualitative component: After completing the six-week intervention, a subset of participants engaged in semi-structured interviews designed to elicit perceptions of how feedback influenced workload, affective responses (e.g., anxiety/frustration), perceived control, and learning regulation (e.g., monitoring and planning).

3.5. Instruments and Measures

Two primary constructs were assessed repeatedly during the intervention phase and subsequently aggregated at the participant level in the analytic plan, with the aggregation described in the manuscript's analysis section.

Perceived cognitive load. Perceived workload was measured using the NASA Task Load Index (NASA-TLX) [21]. A composite score was computed across six dimensions: mental demand, physical demand, temporal demand, effort, performance, and frustration. Internal consistency in this study was high (Cronbach's $\alpha = .88$).

Vocabulary learning strategy use was measured with an adapted Vocabulary Learning Strategies Questionnaire, based on established frameworks [15]. The instrument assessed self-reported frequency of strategy use across memory, cognitive, metacognitive, and social/affective domains using Likert-type ratings. Internal consistency was high (Cronbach's $\alpha = .90$).

Table 2.
Constructs, measures, and reliability

Construct	Instrument	Operationalization	Response format	Reliability (study sample)
Perceived cognitive load	NASA Task Load Index (NASA-TLX)	Composite of six dimensions (mental demand, physical demand, temporal demand, effort, performance, frustration).	Standard NASA-TLX rating format	$\alpha = 0.88$
Vocabulary learning strategy use	Vocabulary Learning Strategies Questionnaire (adapted from Oxford [15])	Composite strategy-use score (memory, cognitive, metacognitive, social/affective strategies)	Likert-type frequency ratings	$\alpha = 0.90$

Note: The parent protocol reports pilot evidence supporting internal consistency for the strategy measure.

4. Data Analysis

4.1. Analytic Approach

Data analysis followed three pre-specified stages. First, descriptive statistics were computed for all focal variables. Internal consistency reliability was evaluated for the NASA Task Load Index (NASA-TLX) workload composite and the adapted Vocabulary Learning Strategies Questionnaire using Cronbach's alpha (α). A Pearson product-moment correlation was then calculated to examine the bivariate association between perceived cognitive load and reported vocabulary learning strategy use.

Because both instruments were administered immediately after each weekly quiz, repeated observations were aggregated at the participant level to support the primary between-condition analyses. Specifically, for each participant, weekly NASA-TLX scores and weekly strategy-use scores were averaged across the six intervention weeks to produce one composite workload score and one composite strategy-use score per participant.

Second, to estimate the effects attributable to the experimental conditions, a path-analytic regression model was fitted in alignment with the conceptual framework. The experimental condition was treated as an exogenous predictor and represented by dummy-coded indicators, with the Control group specified as the reference category. Perceived cognitive load and vocabulary learning strategy use were specified as endogenous outcomes. The principal parameters of interest were the standardized effects (β) of the Immediate Feedback condition on (i) cognitive load and (ii) strategy use. Although the Leaderboards condition was retained to preserve the integrity of the randomized multi-arm design, interpretation in this manuscript centers on isolating the immediate feedback feature. Effects are reported as standardized coefficients (β) with standard errors (SE), *t* statistics, and *p* values. Statistical significance was assessed using two-tailed tests with $\alpha = .05$.

Third, qualitative interview data were used to triangulate and contextualize the quantitative findings. Interview responses were summarized thematically, emphasizing learners' perceptions of (i) workload and affect (e.g., anxiety/frustration), (ii) perceived control, and (iii) self-regulatory behaviors (e.g., monitoring and planning) during weekly quizzing.

4.2. Descriptive Statistics, Reliability, and Correlations

Scale reliability was high for both instruments (NASA-TLX $\alpha = .88$; strategy questionnaire $\alpha = .90$). At the participant level, perceived cognitive load had a mean of 55.2 (SD = 12.3), and vocabulary learning strategy use had a mean of 3.54 (SD = 0.62). The correlation between perceived cognitive load and strategy use was negative ($r = -.30$), indicating that higher perceived workload tended to co-occur with lower reported use of vocabulary learning strategies.

Table 3.Descriptive statistics, reliability, and Pearson correlations (*r*) among primary variables.

Variable	M	SD	α	1	2
1. Cognitive load	55.2	12.3	0.88	—	
2. Vocabulary learning strategies	3.54	0.62	0.90	-0.30	—

Note: Off-diagonal values are Pearson correlation coefficients (*r*). *N* = 60. Higher cognitive- load scores indicate higher perceived workload.

4.3. Effects of Immediate Feedback on Cognitive Load and Strategy Use

Consistent with the hypotheses, immediate feedback demonstrated statistically significant associations with both outcomes relative to the control condition. Participants in the immediate feedback condition reported lower perceived cognitive load ($\beta = -0.42$, $SE = 0.08$, $t = -5.25$, $p < .001$) and higher reported vocabulary learning strategy use ($\beta = 0.55$, $SE = 0.07$, $t = 7.86$, $p < .001$). These results indicate that embedding item-level verification plus brief explanatory feedback within weekly online CSL vocabulary quizzes was linked to reduced perceived workload and greater strategic engagement.

Table 4.Standardized effects of immediate feedback on outcomes (with R^2)

Outcome	Path	β	SE	t	p	Partial R^2
Perceived cognitive load	Immediate feedback \rightarrow Cognitive load	-0.42	0.08	-5.25	< 0.001	0.326
Vocabulary learning strategy use	Immediate feedback \rightarrow Strategy use	0.55	0.07	7.86	< 0.001	0.520

Note: Partial R^2 quantifies the unique proportion of variance in the outcome explained by the Immediate Feedback predictor.

4.4. Qualitative Triangulation

To make the mixed-method triangulation explicit, Table 5 summarizes convergence across three sources: (i) perceived cognitive load (NASA-TLX), (ii) vocabulary learning strategy use questionnaire scores, and (iii) post-intervention semi-structured interviews. Overall, the qualitative themes aligned with the quantitative pattern: immediate feedback reduced perceived workload and supported more strategic vocabulary learning behaviors.

Quantitatively, immediate feedback was associated with lower perceived cognitive load ($\beta = -0.42$, $p < .001$) and higher reported strategy use ($\beta = 0.55$, $p < .001$). In addition, cognitive load and strategy use were negatively correlated ($r = -0.30$), suggesting that higher perceived workload tended to co-occur with lower reported use of vocabulary learning strategies.

Qualitatively, interviewees described immediate feedback as reducing uncertainty and alleviating frustration by clarifying correctness at the moment of response. Participants also reported greater perceived control and confidence during quizzing, and they described more deliberate monitoring and planning behaviors, such as identifying recurring errors, flagging weak vocabulary items for later review, and adjusting study tactics based on feedback received. The triangulation matrix (Table 5) demonstrates where evidence is direct, indirect, or not assessed across the three sources.

Table 5.
Triangulation matrix across three data sources.

Finding / Mechanism	NASA-TLX (quantitative)	Strategy questionnaire (quantitative)	Interviews (qualitative)
H ₁ : Immediate feedback reduces perceived cognitive load	Supported: Immediate feedback → Cognitive load ($\beta = -0.42, p < .001$).	Not directly assessed (outcome differs).	Supported: learners reported reduced uncertainty and lower frustration or anxiety during quizzes.
H ₂ : Immediate feedback increases strategy use	Not directly assessed (outcome differs).	Supported: Immediate feedback → Strategy use ($\beta = 0.55, p < .001$).	Supported: learners described more deliberate monitoring, planning, and targeted review.
Association between cognitive load and strategy use	Supported jointly with strategy measure: $r = -0.30$ (Table 3).	Supported jointly with NASA-TLX: $r = -0.30$ (Table 3).	Convergent narrative: reduced uncertainty/frustration was described as enabling more strategic follow-up (e.g., reviewing weak words).
Mechanism: Uncertainty reduction	Indirectly consistent (lower overall workload).	Indirectly consistent (supports performance monitoring).	Direct support: immediate correctness + brief explanations reduced uncertainty during retrieval practice.
Mechanism: Affective relief (frustration/anxiety)	Indirectly consistent (NASA-TLX composite includes frustration; overall workload reduced).	Not directly assessed.	Direct support: learners reported less frustration/anxiety when feedback clarified errors immediately.
Mechanism: Perceived control/confidence	Indirectly consistent (NASA-TLX includes a perceived performance dimension; overall workload reduced).	Not directly assessed.	Direct support: learners reported greater control and confidence while quizzing.
Mechanism: Monitoring and planning	Not directly assessed.	Direct/partial support (strategy measure includes metacognitive strategy use).	Direct support: learners reported identifying recurring errors, flagging weak items, and adjusting tactics across weeks.

Note: “Supported” indicates direct evidence from that source. “Indirectly consistent” indicates the source does not measure the mechanism directly, but the pattern aligns with the qualitative explanation. “Not directly assessed” indicates no corresponding measure in that source.

5. Discussion

5.1. Immediate Feedback as a Cognitive Load Management Feature

The observed reduction in perceived cognitive load associated with immediate feedback ($\beta = -0.42$) aligns with cognitive load theory, which emphasizes the instructional value of minimizing avoidable sources of extraneous load when learners face high intrinsic task demands. Online CSL vocabulary quizzing requires rapid processing of unfamiliar form–meaning mappings and often orthographic novelty; under these conditions, uncertainty about correctness can inflate perceived workload and divert limited working-memory resources away from learning-relevant processing. Immediate verification paired with brief explanatory information likely mitigated this burden by (i) resolving uncertainty in real time and (ii) streamlining error diagnosis, thereby reducing the need for post hoc reconstruction of mistakes. Learners' interview accounts further support this interpretation, particularly their emphasis on reduced frustration/anxiety and greater perceived control when explanatory feedback was available.

Importantly, the findings suggest that immediate feedback in gamified assessment should not be treated merely as a motivational embellishment. Rather, it appears to function as a cognitively consequential scaffold that can improve the efficiency with which learners navigate timed retrieval practice in online vocabulary tasks.

5.2. Immediate Feedback as a Trigger for Strategic Vocabulary Learning

Immediate feedback was also strongly associated with increased reported strategy use ($\beta=0.55$). This pattern aligns with self-regulated learning accounts in which feedback operates as actionable information supporting monitoring, evaluation, and adaptation. In the present context, prompt correctness signals may have helped learners distinguish known from unknown items, allocate attention more selectively, and adopt more purposeful strategies (e.g., targeted review, deliberate rehearsal, planning subsequent study). The qualitative evidence reinforces this account: learners described using feedback to identify gaps, track weaknesses, and modify study approaches across weeks.

The negative association between cognitive load and strategy use ($r=-.30$) is also theoretically coherent. When perceived workload is elevated, learners may be less able (or less willing) to engage in effortful strategic behaviors; conversely, reductions in perceived workload may create more favorable cognitive and affective conditions for strategy deployment. While the present analyses do not test indirect pathways, the joint pattern suggests that cognitive load management and strategic engagement may be meaningfully intertwined in online CSL vocabulary retrieval practice.

5.3. Theoretical Contributions

This study contributes to the gamified language learning and formative assessment literature by providing feature-level evidence on a ubiquitous but often theoretically under-specified platform mechanism: immediate feedback. Rather than treating gamification as a monolithic “package,” the present design foregrounds item-level verification coupled with brief explanatory feedback as the focal instructional affordance, while preserving the integrity of a randomized multi-arm structure. This framing supports more interpretable theory development because the observed effects can be discussed in relation to a clearly operationalized feature, thereby reducing the common inferential problem in gamification research where multiple concurrent mechanics obscure causal attribution.

The study also advances a process-oriented account of why gamified assessment may matter for demanding L2 vocabulary tasks in online contexts. By centering perceived cognitive load and strategic vocabulary learning, rather than engagement alone, the findings align gamified assessment research more directly with cognitive load theory and self-regulated learning perspectives. The convergent mixed-method pattern further strengthens the plausibility of the proposed mechanisms: immediate explanatory feedback appears to reduce uncertainty and negative affective burden while supporting more deliberate monitoring and planning, positioning feedback not as a superficial motivational embellishment but as a cognitively consequential scaffold in retrieval-based vocabulary practice.

5.4. Practical Implications for Online CSL and Assessment Design

Several actionable implications follow for instructors and instructional designers using Quizizz-like platforms for CSL vocabulary learning and HSK-oriented preparation.

First, immediate feedback should be treated as a core formative design feature for retrieval-based vocabulary quizzing in online CSL contexts, especially when learners face high intrinsic load from unfamiliar characters and rapid form–meaning mapping.

Second, explanatory content matters. Providing only correctness verification may help learners calibrate performance, but brief explanations are more likely to support efficient error diagnosis and reduce uncertainty. At the same time, explanations should be concise and targeted to avoid introducing additional extraneous load.

Third, instructors can improve the instructional yield of feedback by adding “feedback-to-action” affordances that help learners translate item-level feedback into deliberate follow-up. Examples include automatically generating a personalized list of missed words for subsequent review, prompting learners to schedule spaced review sessions, or encouraging brief reflective notes on recurring error types. These affordances operationalize the SRL logic that feedback is most effective when it can be converted into concrete actions.

Finally, because the qualitative findings highlight anxiety/frustration and perceived control, instructors should monitor both cognitive and affective workload during gamified quizzing. In time-pressured formats, immediate clarification can reduce frustration; however, overly dense explanations, excessive speed, or competing gamified elements may, conversely, elevate workload. Balancing clarity, pacing, and informational density is therefore central to cognitively responsible gamified assessment.

5.5. Limitations and Future Research

Several limitations should be considered when interpreting the findings. First, the sample comprised a relatively small group of undergraduates drawn from a single instructional context, which constrains statistical power for detecting smaller effects and limits generalizability across proficiency levels, institutional settings, and learner populations. Replication with larger and more diverse samples, including different CSL proficiency bands and learning environments, would provide a stronger basis for claims about the robustness and boundary conditions of immediate feedback effects in online vocabulary quizzing.

Second, the primary outcomes relied on self-reported perceived workload and self-reported strategy use. While these instruments capture learners' subjective experiences and perceived regulatory behaviors, they do not directly index real-time cognitive processing or observable strategy enactment. Future work should integrate behavioral and trace measures (e.g., response-time dynamics, error trajectories, item revisits, post-quiz review frequency, and learning analytics indicators of regulation), alongside direct vocabulary learning outcomes measured immediately and after a delay. In addition, immediate feedback is not a unitary construct; its effects may vary with timing, informativeness, and explanatory granularity. More fine-grained experiments comparing verification-only versus elaborated feedback, immediate versus delayed feedback, and generic versus adaptive explanations would clarify which configurations best optimize cognitive efficiency and strategic engagement. Finally, learner characteristics such as anxiety, self-efficacy, competitiveness, and baseline self-regulation may moderate feedback use, and hypothesized pathways may involve mediation through uncertainty reduction or affective relief; future studies should therefore test moderated and mediated models to specify for whom, how, and under what conditions immediate feedback yields the greatest benefits.

6. Conclusion and Recommendations

6.1. Conclusion

This study examined the role of immediate feedback in gamified online Chinese vocabulary quizzing during a six-week HSK preparation course. The results indicate that, compared with a non-gamified control condition, embedding item-level verification with brief explanatory feedback is associated with (a) lower perceived cognitive load ($\beta = -0.42$) and (b) higher reported use of vocabulary learning strategies ($\beta = 0.55$). Qualitative interview themes converge with these quantitative patterns, suggesting that immediate feedback reduces uncertainty, lowers anxiety/frustration, strengthens perceived control, and supports more deliberate monitoring and planning during weekly vocabulary retrieval practice. Collectively, the findings position immediate feedback as a cognitively meaningful scaffold in online CSL quizzing, one that appears to shape learners' experienced workload and their strategic approach to vocabulary learning.

6.2. Recommendations for Instructional Practice

To enhance the pedagogical value of gamified vocabulary quizzing in online Chinese-as-a-second-language (CSL) contexts, instructors and program designers should treat immediate feedback as a default formative design feature rather than an optional add-on, particularly when learners are required to process unfamiliar characters and dense lexical sets that impose high intrinsic cognitive demands. In practice, this entails ensuring that quizzes provide not only correctness verification but also brief, actionable explanatory cues that enable rapid error diagnosis while avoiding excessive informational

detail that could increase extraneous load. Explanations should be tightly aligned with common form, meaning confusions and presented concisely to support learners' moment-to-moment calibration during retrieval practice, thereby promoting cognitive efficiency without disrupting quiz flow.

In addition, instructional implementations should be designed so that feedback reliably translates into follow-up learning activities. Platforms or course procedures can operationalize "feedback-to-action" pathways by automatically compiling individualized review lists of missed items, prompting learners to schedule spaced review opportunities, and encouraging short, structured reflections on recurrent error types to strengthen metacognitive monitoring and planning. Because learners' experience of gamified quizzing is shaped by both cognitive and affective workload, instructors should also monitor perceived task demands and emotional responses, especially frustration and anxiety, and adjust pacing variables such as time limits, quiz length, and feedback density to ensure that feedback functions as a scaffold that reduces unnecessary burden rather than becoming an additional source of overload.

6.3. Recommendations for Research

Future research should extend the present evidence base through systematic replication and methodological enrichment. Studies conducted across multiple sites, proficiency levels, and course formats would help establish generalizability and clarify whether the observed associations hold under different instructional constraints, lexical difficulty profiles, and cultural contexts. In parallel, incorporating objective indicators of learning processes and outcomes, such as learning analytics, response-time patterns, error correction trajectories, and delayed retention measures, would allow researchers to test whether reductions in perceived workload and increases in reported strategy use translate into durable vocabulary gains and transfer.

Transparency:

The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.

Copyright:

© 2026 by the authors. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

References

- [1] S. Deterding, D. Dixon, R. Khaled, and L. Nacke, "From game design elements to gamefulness: defining" gamification", in *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, 2011.
- [2] K. M. Kapp, *The gamification of learning and instruction: Game-based methods and strategies for training and education*. Hoboken, NJ: John Wiley & Sons, 2012.
- [3] K. Seaborn and D. I. Fels, "Gamification in theory and action: A survey," *International Journal of human-computer studies*, vol. 74, pp. 14–31, 2015. <https://doi.org/10.1016/j.ijhcs.2014.09.006>
- [4] H. L. Roediger III and J. D. Karpicke, "Test-enhanced learning: Taking memory tests improves long-term retention," *Psychological Science*, vol. 17, no. 3, pp. 249–255, 2006. <https://doi.org/10.1111/j.1467-9280.2006.01693.x>
- [5] J. D. Karpicke and H. L. Roediger III, "The critical importance of retrieval for learning," *Science*, vol. 319, no. 5865, pp. 966–968, 2008. <https://doi.org/10.1126/science.1152408>
- [6] Z. Zainuddin, M. Shujahat, and H. Haruna, "Gamified assessment in higher education: A systematic review," *Interactive Learning Environments*, vol. 28, no. 2, pp. 147–165, 2020.
- [7] J. Hattie and H. Timperley, "The power of feedback," *Review of Educational Research*, vol. 77, no. 1, pp. 81–112, 2007. <https://doi.org/10.3102/003465430298487>
- [8] V. J. Shute, "Focus on formative feedback," *Review of Educational Research*, vol. 78, no. 1, pp. 153–189, 2008. <https://doi.org/10.3102/0034654307313795>

- [9] R. L. Bangert-Drowns, C.-L. C. Kulik, J. A. Kulik, and M. Morgan, "The instructional effect of feedback in test-like events," *Review of Educational Research*, vol. 61, no. 2, pp. 213-238, 1991. <https://doi.org/10.3102/00346543061002213>
- [10] A. N. Kluger and A. DeNisi, "The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory," *Psychological Bulletin*, vol. 119, no. 2, pp. 254-284, 1996.
- [11] F. Paas, A. Renkl, and J. Sweller, "Cognitive load theory and instructional design: Recent developments," *Educational Psychologist*, vol. 38, no. 1, pp. 1-4, 2003. https://doi.org/10.1207/S15326985EP3801_1
- [12] J. J. Van Merriënboer and J. Sweller, "Cognitive load theory in health professional education: Design principles and strategies," *Medical Education*, vol. 44, no. 1, pp. 85-93, 2010. <https://doi.org/10.1111/j.1365-2923.2009.03498.x>
- [13] J. Sweller, "Cognitive load theory and educational technology," *Educational Technology Research and Development*, vol. 68, no. 1, pp. 1-16, 2020. <https://doi.org/10.1007/s11423-019-09701-3>
- [14] L. Ding, E. Er, and M. Orey, "Impact of gamification on learners' cognitive load," *Computers & Education*, p. 18, 2022.
- [15] R. L. Oxford, *Language learning strategies: What every teacher should know*. Boston, MA: Heinle & Heinle Publishers, 1990.
- [16] Y. Gu and R. K. Johnson, "Vocabulary learning strategies and language learning outcomes," *Language Learning*, vol. 46, no. 4, pp. 643-679, 1996. <https://doi.org/10.1111/j.1467-1770.1996.tb01355.x>Digital
- [17] N. Schmitt, *Vocabulary learning strategies*. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy*. Cambridge, UK: Cambridge University Press, 1997.
- [18] I. S. P. Nation, *Learning vocabulary in another language*. Cambridge, UK: Cambridge University Press, 2001.
- [19] D. L. Butler and P. H. Winne, "Feedback and self-regulated learning: A theoretical synthesis," *Review of Educational Research*, vol. 65, no. 3, pp. 245-281, 1995. <https://doi.org/10.3102/00346543065003245>
- [20] D. J. Nicol and D. Macfarlane-Dick, "Formative assessment and self-regulated learning: A model and seven principles of good feedback practice," *Studies in Higher Education*, vol. 31, no. 2, pp. 199-218, 2006. <https://doi.org/10.1080/03075070600572090>
- [21] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research," *Advances in Psychology*, vol. 52, pp. 139-183, 1988. [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)