Jurnal Pembelajaran Bahasa dan Sastra
Volume 5, Nomor 6, November 2026
DOI: https://doi.org/10.55909/jpbs.v5i6.1910

https://pembahas.dialeks.id/index.php/jp
p-ISSN : 2809-5367
e-ISSN : 2809-6355

Needs Analysis for a Context-Based HSK 4 Simulation Website as a Preparation Tool for the Mandarin Language Proficiency Exam

Icha Fais Nurul Karimah[1*], Aprilia Ruby Wikarti[2], Rizky Wardhani[3], Anisya Putri Nasution[4], Khofifah Nur Asadiyah[5], Siti Fauziyah[6]

[123456]Prodi Pendidikan Bahasa Mandarin, Universitas Negeri Jakarta, DKI Jakarta, Indonesia

*E-mail: icha.fais@unj.ac.id

ABSTRACT

This study aims to examine user requirements for a contextual-based Hanyu Shuiping Kaoshi Level 4 (HSK 4) simulation *website*, serving as the foundational step for developing effective Mandarin language test preparation media. Conducted as the needs analysis phase within a broader Research and Development (R&D) framework, this research took place among Mandarin instructors between January and February 2026. Primary data were gathered through questionnaires distributed to 50 Mandarin teachers and instructors who participated as respondents. The research instruments explored various dimensions, including prior HSK experience, utilization of practice platforms, common constraints of existing media, required features, scoring and feedback mechanisms, accessibility, simulation attributes, contextual elements, language components, and device preferences. Quantitative data from closed-ended questions were analyzed descriptively using frequencies, percentages, and mean scores, whereas qualitative insights from open-ended responses were synthesized based on prevailing answer tendencies. The findings reveal that while 80% of respondents have previously relied on digital websites or applications for HSK preparation, the availability of comprehensive yet free resources remains only moderate, averaging 3.12 out of a 5-point scale. Primary obstacles identified by users involve restricted free practice options (72%), paywalled features (68%), the lack of full-length mock exams (44%), and the absence of detailed answer explanations (42%). Notably, all respondents (100%) expressed a strong demand for complete HSK simulations, followed by listening and writing drills at 84% each, reading exercises at 82%, and integrated timers at 80%. Furthermore, participants highlighted the necessity for automated grading, clear solution walkthroughs, cross-device compatibility, and materials embedded within everyday life contexts. Consequently, it is evident that an HSK 4 simulation website must be engineered as an all-in-one practice ecosystem that successfully bridges exam authenticity, access flexibility, constructive feedback, progress tracking, and contextualized content delivery.

Keywords: needs analysis, context, simulation website, tool, Mandarin language

Analisis Kebutuhan Website Simulasi HSK 4 Berbasis Kontekstual sebagai Media Persiapan Ujian Bahasa Mandarin

Penelitian ini bertujuan untuk mengkaji kebutuhan pengguna terhadap pengembangan website simulasi Hanyu Shuiping Kaoshi tingkat 4 (HSK 4) berbasis kontekstual sebagai landasan utama dalam menyediakan media persiapan ujian bahasa Mandarin yang efektif. Penelitian ini dilaksanakan di lingkungan para pengajar bahasa Mandarin pada bulan Maret hingga Juni 2026. Data primer dikumpulkan melalui penyebaran angket kepada 50 orang guru dan instruktur bahasa Mandarin yang bertindak sebagai responden utama. Analisis data tertutup dilakukan secara deskriptif dengan menghitung frekuensi, persentase, serta nilai rata-rata. Sementara data dari tanggapan terbuka dikelompokkan berdasarkan pola jawaban yang dominan. Hasil temuan memperlihatkan bahwa meskipun 80% responden sudah terbiasa memanfaatkan platform digital untuk belajar HSK, ketersediaan media yang bersifat gratis sekaligus menyeluruh masih dinilai berada pada kategori sedang dengan skor rata-rata 3,12 dari skala maksimal 5. Sejumlah kendala utama yang dikeluhkan meliputi terbatasnya porsi latihan gratis (72%), keharusan membayar untuk fitur lanjutan (68%), kurangnya simulasi ujian yang utuh (44%), serta ketiadaan pembahasan jawaban secara mendetail (42%). Menariknya, seluruh responden (100%) menyatakan butuh terhadap simulasi HSK yang komprehensif, disusul oleh kebutuhan akan latihan menyimak serta menulis masing-masing sebesar 84%, latihan membaca 82%, serta fitur pengatur waktu sebesar 80%. Selain itu, para responden juga mengharapkan adanya sistem penilaian otomatis, pembahasan soal yang jelas, kemudahan akses lintas perangkat, serta materi yang dikemas dekat dengan situasi kehidupan sehari-hari. Berdasarkan paparan tersebut, dapat disimpulkan bahwa perancangan website simulasi HSK 4 ini sangat mendesak untuk diwujudkan sebagai sarana latihan terpadu yang mampu memadukan keaslian format ujian, fleksibilitas akses, umpan balik yang membangun, fitur pantauan perkembangan belajar, serta penyajian konten yang kontekstual.

Kata kunci: analisis kebutuhan, website simulasi, kontekstual, media persiapan ujian, bahasa Mandarin

Submitted	Accepted	Published
27/10/2026	11/11/2026	24/11/2026

Citation	Karimah, I. F. N., Wikarti, A. R., Wardhani, R., Nasution, A. P., Asadiyah, K. N, & Fauziyah, S. (2026). Needs Analysis for a Context-Based HSK 4 Simulation Website as a Preparation Tool for the Chinese Language Proficiency Exam. *Jurnal Pembelajaran Bahasa dan Sastra, Volume 5, Nomor 6, November 2026, 5345-5356*. DOI: https://doi.org/10.55909/jpbs.v5i6.1910

Publisher
Raja Zulkarnain Education Foundation

https://pembahas.dialeks.id/index.php/jp
p-ISSN : 2809-5367
e-ISSN : 2809-6355
Jurnal Pembelajaran Bahasa dan Sastra
Volume 5, Nomor 6, November 2026
DOI: https://doi.org/10.55909/jpbs.v5i6.1910

INTRODUCTION

Facing *the Hanyu Shuiping Kaoshi* Level 4 (HSK 4) exam requires Mandarin learners to be able to master listening, reading, and writing skills under strict time constraints (Azizah et al., 2025). In practice, most learners still rely on conventional printed books or static question banks that are not yet able to fully represent the actual exam experience (Su, 2015:45–52). The lack of variety in digital-based exercises, the absence of in-depth answer reviews, and materials often isolated from everyday communication are classic obstacles frequently complained about. These various obstacles in the field underscore the urgency of providing more adaptive preparation media that directly rely on the real needs of users.

The transformation of educational technology has also opened up space for the development of web-based assessment (*computer-based language assessment*). The advantages of digital systems allow the integration of audio elements, timer features, automatic score recording, and instant corrective responses within a unified training ecosystem (Douglas & Hegelheimer, 2007:115–132). From a pedagogical perspective, Chapelle et al. (2003:409–439) emphasize that the design of computer-based assessments must be based on functional validity and meaningful interaction for users. Furthermore, the flexibility of digital platforms has proven capable of supporting participants' *self-directed learning* outside of formal lecture hours (Lai et al., 2016:40–60).

Nevertheless, the effectiveness of a learning technology product is highly determined by the accuracy of its planning through needs *analysis*. Cowling (2007:426–442) views needs analysis as an essential step to map the gap between empirical field conditions and ideal media specifications. Especially in the field of HSK certification, Mo et al. (2022:169–188) emphasize that identifying expectations and complaints from prospective users is an absolute foundation so that the developed features have strong pedagogical relevance. The washback effect of teaching a test instrument also demands contextual material presentation so that learners not only memorize language patterns but also understand their communicative functions (Li & Hu, 2025).

Several previous studies have extensively explored the realm of assessment and language learning media. Research by Azizah et al. (2025), for example, highlights that understanding dialogue contexts remains the biggest stumbling block for students in HSK 4 simulations. Similar findings were revealed by Bahari (2021:24–38), which proves that *computer-mediated feedback* plays a crucial role in sharpening students' self-evaluation. On the other hand, García Botero et al. (2019:71–97) noted that optimizing gadgets in cross-platform learning can expand practice access, as long as supported by structured feature layouts. Nevertheless, research gaps remain wide open because not many developers have mapped out the specific needs for web-based HSK 4 simulation platforms that comprehensively combine integrated assessment with contextual content.

Based on the above explanation, the research problem formulation is formulated as follows:

1) What is the profile of the experiences and challenges faced by prospective users in utilizing the current HSK 4 preparation media?
2) What are the specifications of prospective users' needs regarding features, assessment systems, feedback, and contextual presentation in designing *the HSK 4 simulation* website?

In line with the above problem formulation, this study aims to analyze the needs of prospective users for the contextual-based HSK 4 simulation website, covering aspects of exam experience, existing media conditions, technical and pedagogical constraints, feature requirements, assessment and feedback systems, access flexibility, simulation characteristics, contextual content, as well as language and device preferences.

https://pembahas.dialeks.id/index.php/jp
p-ISSN : 2809-5367
e-ISSN : 2809-6355

The benefits of conducting this research can be mapped into two domains. Practically, the results of this study are expected to serve as empirical guidelines for software developers and educators in designing targeted HSK 4 training media. Theoretically, this research is expected to enrich the discourse on *needs analysis* in the development of web-based language assessment (*R&D*) media, especially within the scope of Mandarin language learning.

METHOD

This study adopts a descriptive needs analysis design placed as the initial phase within the research and development (*R*&D) framework based on the Sugiyono (2020) model. The research first focuses on the needs *analysis* stage to dissect field dynamics, map prospective user profiles, and formulate feature priority scales before the HSK 4 simulation website architecture is executed (Brown, 2016). This descriptive orientation was chosen so that exploration can present a comprehensive presentation of quantitative and qualitative data regarding the empirical context of the target audience, without stepping into hypothesis testing the impact of media interventions (Creswell & Creswell, 2023).

The data collection process took place over two months, from January to February 2026. The research was set within the academic ecosystem of Mandarin language teachers and instructors in Indonesia, connected through professional community networks and online-based teacher forums. The participants involved included 50 Mandarin language teachers from various educational institutions, ranging from public high schools, private schools, vocational schools, to language guidance institutions. In terms of service period, 36 respondents (72%) were recorded to have less than one year of teaching experience, 5 respondents (10%) were in the range of 1–5 years and 6–10 years, while 4 respondents (8%) had been practicing their profession for more than a decade. From the exam track record perspective, 33 participants (66%) admitted they had never sat directly in the HSK test seat, while the remaining 17 (34%) had experience as exam participants.

The research implementation flow is designed through four systematic stages: (1) initial exploratory study and questionnaire instrument preparation; (2) process of reviewing and refining the instrument's content; (3) distributing questionnaires in a distributed manner to target participants through digital media; and (4) data processing, thematic analysis, and explanation of required outputs into functional product specifications.

Data collection relied on a needs analysis questionnaire instrument distributed centrally using the *Google Forms platform*. The questionnaire construct summarizes a combination of closed and open-ended question items. The closed section covers the following variables: (1) professional background and history of HSK; (2) exam participation orientation; (3) factors inhibiting exam participation; (4) utilization of current learning facilities; (5) constraints on existing material sources; (6) perception of the urgency of supporting media; (7) level of expectation regarding simulation, skill-based drills, automatic scoring, and progress tracking; (8) device flexibility and language of instruction; and (9) preferences for presenting contextual material. Some closed items were measured using a five-level Likert scale (1 = Strongly Disagree/Strongly Not Needed up to 5 = Strongly Agree/Very Necessary).

Meanwhile, the questionnaire section ended with five open-ended questions to gather qualitative input related to the most crucial challenges, relevant communication context topics, and architectural input for platform development. The quantitative data processing technique was carried out through descriptive analysis to obtain an overview of frequency, percentage distribution, as well as the achievement of mean values and *standard deviation* for each indicator. All descriptive statistical computations were processed using SPSS (statistical*package for the Social Sciences*) software to ensure the accuracy and

Jurnal Pembelajaran Bahasa dan Sastra
Volume 5, Nomor 6, November 2026
DOI: https://doi.org/10.55909/jpbs.v5i6.1910

precision of the numerical calculations. Frequency and percentage values were used to map demographic trends and identify constraints, while *the calculation of mean* values became the main pillar in determining the criteria for the level of feature needs. The descriptive responses to the open-ended questions were analyzed through content *analysis* through information reduction, thematic pattern coding, and qualitative inference. All data analysis in this phase was exploratory-descriptive oriented without involving inferential statistics because the media products had not yet entered the effectiveness testing stage.

RESULTS

1. Challenges for Mandarin Language Teachers in Preparing for the HSK Exam

Based on the processing of questionnaire data, the challenges faced by Mandarin language teachers in preparing for the HSK exam include aspects of personal experience, financial constraints, and the limitations of currently available learning resources. The detailed findings in this section are formulated as follows:

a. There is a significant gap between the level of professional awareness of HSK and direct experience of taking the exam, where the majority of respondents (72%, n = 36) do not yet have an HSK certificate and 66% (n = 33) have never taken the HSK test in person, although 96% (n = 48) acknowledge HSK as proof of teaching competence.
b. The relatively high exam registration fee is the main obstacle factor for participation (64%, n=32), followed by low self-confidence (48%, n=24), limited exam schedules (34%, n=17), and lack of understanding of the HSK exam structure (30%, n=15).
c. Existing learning resources accessed by teachers are still felt to not yet meet optimal preparation needs (M = 3.20; SD = 0.83), and the ease of finding complete free training websites falls into the moderate category (M = 3.10; SD = 0.93).
d. Limited free practice (72%, n=36) and dominance of paid features (68%, n =34) are the main obstacles in existing digital media, followed by the absence of a complete exam simulation (44%, n=22), absence of answer discussions (42%, n =21), and the unavailability of a feature to monitor the progress of practice results (38%, n=19).

Table-1
Characteristics of Teaching Experience and HSK Participants' Experiences

1 year of teaching experience <	36	72
1–5 years teaching experience	5	10
6–10 years of teaching experience	5	10
10 years of teaching experience >	4	8
Does not yet have an HSK certificate	36	72
Possess an HSK certificate	14	28
Never participated in HSK	33	66
Have participated in HSK	17	34
To understand HSK 4 as a reference in the research context	44	88
Recognizing HSK as proof of competence	48	96

2. Barriers to Taking or Preparing for the HSK Exam

When asked to explain the reasons why they have not taken or rarely took HSK, respondents mostly mentioned relatively high exam fees (64%) and lack of confidence in their own abilities (48%). The next obstacles are limited exam schedules (34%), lack of understanding of HSK exam structure (30%), limited study time (26%), and exam locations that are difficult to reach (20%). Some respondents also mentioned not feeling the need for certificates (18%), difficulty finding free HSK practice (16%), and a complicated

https://pembahas.dialeks.id/index.php/jp
p-ISSN : 2809-5367
e-ISSN : 2809-6355

registration process (12%). This pattern shows that the decision to participate in HSK is influenced not only by language readiness but also by access, information, time, and self-confidence.

Table-2
The Reason Respondents Have Not or Rarely Take the HSK Exam

Reasons	n	%
Relatively expensive exam fees	32	64
Lack of confidence in abilities	24	48
Limited exam schedule	17	34
Not yet understanding the structure of the HSK exam	15	30
Less time to study	13	26
Exam locations are hard to reach	10	20
Not yet feeling the need for the HSK certificate	9	18
It's hard to find free HSK training	8	16
The registration process is considered complicated	6	12

3. Preparation Practices and the Limitations of Available Learning Resources

The use of digital resources is quite common among respondents. As many as 40 respondents (80%) have used *websites* or applications to study HSK. The most commonly used media for preparation are HSK exercise books (68%), HSK practice websites (66%), and learning applications (54%). To obtain questions, 72% use free websites and 62% download questions from the internet. Nevertheless, the ease of finding *free* and complete HSK simulation websites only obtained an average *of M* = 3.10 (*SD* = 0.93), while the assessment that the available learning resources met the needs was at an average *of M* = 3.20 (*SD* = 0.83). The most frequently reported obstacles were the limitations of free training and paid features. Respondents also still faced incomplete simulations, minimal discussion, limited monitoring functions, and absence of automatic assessment.

Table-3
Obstacles to the Available HSK Learning Resources

Relatively expensive exam fees	32	64
Lack of confidence in abilities	24	48
Limited exam schedule	17	34
Not yet understanding the structure of the HSK exam	15	30
Less time to study	13	26
Exam locations are hard to reach	10	20
Not yet feeling the need for the HSK certificate	9	18
It's hard to find free HSK training	8	16
The registration process is considered complicated	6	12

4. The Need for Flexible and Independent Preparation

The need for flexible and independently usable media consistently appears on the five-point scale. The statement that Mandarin language teachers need additional media to prepare for HSK received *M* = 4.56 (*SD* = 0.64), with 92% of respondents choosing to agree or strongly agree. The need for easy-to-use practice media for learners obtained *M* = 4.60 (*SD* = 0.53; 98% agree/ strongly agree). In terms of flexibility, exercises accessible anytime obtained *M* = 4.72 (*SD* = 0.54; 96%), while the need for media to independently measure ability obtained *M* = 4.74 (*SD* = 0.49; 98%). Interestingly, the statement that *the simulation website* can improve readiness for HSK obtained a more moderate average, namely *M* = 3.68 (*SD* = 0.47; 68% agree). This difference indicates that respondents are very clear about the functions they need, but the benefits of increased readiness cannot yet be assumed without further testing.

5. Priority Functions for Web-Based HSK Support

The strongest priority is on the exam simulation function. All respondents chose the complete HSK simulation as the feature that needs to be available. Listening and writing exercises were chosen by 84%, reading exercises by 82%, timer and grammar materials 80%, and audio exercises 78%. Vocabulary lists, free access without subscription fees, answer keys and discussions, and automatic scoring were each chosen by 76%. When respondents were asked to choose the most needed features, the order became more selective: complete exam simulation selected 90%, answer key and discussion 88%, vocabulary and grammar material 78%, practice audio 64%, automatic assessment 62%, and score history and progress 56%. Thus, simulation, discussion, supporting materials, and feedback emerged as a consistent core function group for the two types of questions.

Table-4
The most popular web-based HSK feature

Features	n	%
Complete HSK simulation	50	100
Listening exercises	42	84
Writing exercises	42	84
Reading exercises	41	82
Exam timer	40	80
HSK grammar material	40	80
Exercise audio	39	78
Daftar kosakata HSK	38	76
Free access with no subscription fee	38	76
Answer key and discussion	38	76
Automatic assessment	38	76
Graph of training progress	35	70
Exam simulation mode	34	68
User value history	33	66
Training mode	32	64
Topic-based exercises	30	60
Recommended materials	30	60
Smartphone-friendly display	29	58

6. Feedback, Access, and Material Presentation Preferences

Scale items reinforce the importance of feedback and simulation experiences. The need to assess performance in each exam section obtained $M = 4.84$ ($SD = 0.42$), automatic scores $M = 4.74$ ($SD = 0.49$), answer keys and explanations $M = 4.80$ ($SD = 0.45$), simulation history $M = 4.78$ ($SD = 0.51$), and development information $M = 4.80$ ($SD = 0.45$). The simulation resembling an exam obtained $M = 4.68$ ($SD = 0.55$; 96% agree/strongly agree), while the needs for listening, reading, and writing reached $M = 4.80$ ($SD = 0.45$). Audio clarity received the highest scores along with several access aspects, namely $M = 4.84$ ($SD = 0.37$; 100% agree/strongly agree), while exam timer obtained $M = 4.76$ ($SD = 0.52$). From the access side, usage via computer and smartphone, as well as simple interfaces, each obtained $M = 4.84$ ($SD = 0.42$). For the device, smartphone, and laptop questions were chosen by 74% of respondents respectively. As many as 74% also chose a combination of Indonesian, Mandarin, and English as their preferred language of instruction. The preference for presenting material close to the language usage situation is also high: linking questions to everyday life obtained $M = 4.60$ ($SD = 0.57$), vocabulary presentation through conversations or real situations $M = 4.70$ ($SD = 0.51$), and using context to help vocabulary understanding $M = 4.80$ ($SD = 0.45$). These results place situational elements as preference, not as claims that all test items need to be transformed into contextual material.

7. Strengthening Findings from the Open Response

The five open-ended questions showed patterns consistent with the quantitative results. Recurring difficulties related to limited study time, exam costs or feature access, difficulty listening to audio at natural speed, mastery of vocabulary and grammar, Hanzi writing, and the absence of structured preparation guides. When evaluating *websites* or applications they had used, respondents

https://pembahas.dialeks.id/index.php/jp
p-ISSN : 2809-5367
e-ISSN : 2809-6355

again mentioned limited free questions, paid features, lack of adequate discussions, poor audio, interfaces that did not resemble exams, and the absence of development analysis. Development suggestions were also consistent: respondents wanted a broader question bank, discussion of each answer, timers, simulations closer to test procedures, progress tracking, practice recommendations based on weaknesses, and convenient access via laptops and smartphones. This open response reinforces the interpretation that user needs are integrated and do not stop at increasing the number of questions.

DISCUSSION

The main argument in this study shows that preparation for the HSK Exam (Level 4) for teachers and prospective users cannot be reduced merely to a matter of language proficiency. The needs analysis results show that professional awareness of HSK certification is very high, but hindered by various structural, psychological, and accessibility factors. The following section outlines the theoretical and empirical arguments related to these findings, their implications for development design, as well as the strengths and limitations of the research.

The Relevance of HSK Professionals and Participation Barriers

The research findings show a gap between respondents' awareness of the importance of HSK certification and their direct experience as test participants. Although 96% of respondents recognize HSK as proof of instructor competence and 88% understand its use as an academic reference, the majority have never taken the exam in person. These obstacles are dominated by cost factors, self-doubt, time constraints, and lack of understanding of exam structure. This phenomenon broadens understanding of *the concept of washback* in language assessment. Li and Hu (2025) emphasize that test characteristics, question formats, and assessment criteria can trigger the impact of *washback* on participants' learning strategies. On the other hand, Yang and Jirawit (2025) place familiarization of the test format and preparation barriers as crucial variables in testing experience.

From the perspective of assessment literacy, these findings align with the views of Estaji (2024) and Yu and Xu (2024), who state that understanding assessment influences teacher agency and preparation strategies in the digital environment. Mandarin teachers not only act as classroom facilitators but also subjects who need clarity of exam information. Therefore, web-based intervention platforms need to focus on bridging the gap in test structure recognition, not just increasing the pile of practice questions.

Sources Available, but Still Fragmented

Analysis of existing learning resources shows that the main problem is not the absence of digital materials, but the fragmentation of functions and the absence of integration of learning flows. Although most respondents have utilized books, free websites, or applications, the ease of finding complete and free simulation platforms is still considered moderate. Users often have to collect vocabulary materials from one source, audio files from another source, and practice questions separately without an ecosystem that unifies simulations, feedback, and progress tracking.

This condition is relevant to the principles *of Computer-Assisted Language Learning* (CALL). Aljohani (2025) emphasizes the importance of autonomy, practicality, and suitability of digital tasks in supporting learner achievement, while Morgana (2024) highlights that technological effectiveness heavily depends on feature design that directly targets exam subskills. At the local level, studies by Sartika et al. (2024) and Ventivani et al. (2026) confirm that digital media integration in Indonesia has become a real necessity in Mandarin language learning, although design quality and accessibility challenges remain major obstacles.

Jurnal Pembelajaran Bahasa dan Sastra
Volume 5, Nomor 6, November 2026
DOI: https://doi.org/10.55909/jpbs.v5i6.1910

Simulation as the Top Design Priority

Based on respondent preferences, the complete HSK simulation feature takes absolute priority, followed by separate practice modules, answer discussions, automatic assessment, and progress track record. This pattern supports a clear separation between practice *mode* and timed simulation mode. Practice mode provides room for exploration, improvement, and deep understanding, while simulation mode maintains the rigor of real exam procedures. However, strong support for simulation should not be interpreted as a claim of instant effectiveness. A moderate value in the statement that a simulation website automatically increases readiness indicates that functional needs and product effectiveness levels are two different empirical domains.

Referring to the views of Bachman and Palmer (2010) and Fulcher (2010), the interpretation of assessment results and the use of practice scores must be supported by adequate evidence of validity. Therefore, the scores produced by the simulation platform in this early phase are positioned as self-practice metrics, not as an absolute predictor of official exam passing.

Pass Covers and Monitoring Extend Preparation Functions Beyond Scoring

The high expectations for automatic feedback, detailed scores per skill, answer keys with discussions, and progress graphs indicate that users need meaningful formative information after completing tasks. This aligns with the digital formative assessment model proposed by Sawaki et al. (2024), as well as the emphasis by Shi and Aryadoust (2024) on the need for strict validation of the quality of automatic feedback.

Furthermore, the functions of history tracking and answer review support the *self-regulated learning* process. As shown by Zhang and Yu (2024) and Zhang, Liu, and Wilson (2024), self-engagement in digital platforms demands a navigation structure that guides users to systematically move from the directed training stage to full simulation. Usability aspects, *such as simple interfaces and cross-device compatibility (computers and smartphones), also reinforce the findings of Maahs et al. (2025) and Lee and Jeon (2024), that ease of access is an integral part of the pedagogical effectiveness of technology.*

Cross-Device Access and Material Presentation as Supporting Needs

The need for device flexibility and multilingual interface support indicates that operational comfort is the main pillar of successful technology adoption. On the content side, the preference for presenting material based on real situations (everyday life contexts) serves as a supporting strategy to facilitate vocabulary and sentence structure understanding in practice mode. However, the application of this context is limited to avoid altering the core construct of the official HSK *blueprint*, especially when users enter formal exam simulation mode.

From Evidence of Needs to Design Requirements and Evaluation of Research Limitations

The entire range of findings, from financial and psychological barriers, fragmentation of learning resources, to priority of interactive features, has been systematically explained in the platform's functional requirements draft. This mapping serves as the rational foundation for developing evidence-based products.

Table-5
Implications of evidence-based design for further platform development

Proof of Need	Interpretation	Design Implications
64% cost; 48% lack confidence; 30% exam structure	Readiness is influenced by access, confidence, & familiarization	Low-cost access & test structure guidance
100% complete simulation; 90% top priority	High demand for procedural familiarization	Complete simulation mode with *timer*

https://pembahas.dialeks.id/index.php/jp
p-ISSN : 2809-5367
e-ISSN : 2809-6355

84% listening/writing; 82% reading	Directed subskill training needs	Separate training modules based on skills
76% automatic assessment; M = 4.84 results per section	Direct performance feedback needs	Practice scores by section and summary of achievements
76% of discussions; 88% of most needed features	The need to understand the reasons for right or wrong	Facilities for reviewing answers and in-depth discussions
70% progress graph; 66% grade history	The need to monitor periodic developments	Trial history and progress graph visualization
M = 4.84 across devices; 74% *smartphone*, 74% laptop	The need for flexibility in self-use	Responsive design (*cross-device*) and simple navigation
M = 4.60–4.80 presentation based on real situations	The need for contextualization of training materials	Integration of situational context in training modules

Although the results of this study have provided an overview of user needs and *platform* development directions, there are still several things that need to be considered in the next stages. The findings in this study mainly describe features considered important by respondents. Thus, these needs cannot yet be directly interpreted as evidence that all proposed features will be easy to use or effective in assisting HSK 4 exam preparation. This needs to be further studied through direct product development and testing.

Further *platform* development can begin by preparing prototypes based on identified priority needs. *The prototype* needs to be reviewed by Mandarin language experts and learning technology experts to ensure the suitability of the materials, training formats, and available functions. After that, testing with prospective users is necessary to determine whether the usage flow is easy to understand, whether the instructions are clear enough, and whether the feedback provided helps users understand their mistakes. The test results can be used as a basis to improve features before the platform is used more widely.

The assessment aspect also requires special attention. Each skill in HSK 4 has different characteristics, so not all can be assessed in the same way. Automatic scores, for example, can be used for certain types of questions, but writing skill assessment requires broader consideration. Therefore, the development of assessment features needs to be adjusted to the characteristics of the questions and the purpose of the exercise. In addition, simulation materials must refer to the exam format that forms the basis of development, while still considering the source material and copyright provisions.

Findings regarding the need for progress charts and grade history also show that users not only need a place to practice but also a means to monitor their learning progress. This information should be presented simply so that users can recognize areas they have mastered and skills that still need improvement. In this way, the progress monitoring feature not only functions as a record of practice results but can also help users determine their next learning steps.

Based on these considerations, the next research stage is directed towards developing a contextual-based HSK 4 simulation platform prototype, followed by user testing. The initial evaluation can focus on functional feasibility, ease of use, material suitability, and the benefits of feedback for users. If the platform meets these criteria, further research can assess its contribution to participants' readiness to face the HSK 4 exam through appropriate evaluation design. Thus, product development is carried out gradually, based on the needs found in this study while also considering the test results at each stage.

CONCLUSION

This study examines the preparation needs of 50 Mandarin teachers and instructors, focusing on barriers to taking tests, limited available learning resources, and forms of web-based support considered important. The analysis results show that awareness of the professional relevance of HSK is much higher compared to

direct experience taking tests. The main barriers to participation are factors such as cost, lack of self-confidence, limited exam schedules, limited understanding of test structure, and limited study time. In terms of learning resources, respondents reported that although digital media has been widely used, the main challenges lie in the limited free practice, dominance of paid features, lack of full simulations, minimal discussions, and insufficient facilities for monitoring progress. In line with this, the highest level of demand is directed towards complete simulation features, practice by subskill, answer keys with in-depth discussions, timers, automatic assessments, progress tracks, and easy cross-device accessibility.

From a scientific perspective, this study expands the discourse in the field of *Computer-Assisted Language Learning* (CALL) and language assessment by positioning Mandarin language teachers as a stakeholder group with specific characteristics and needs. The main contribution of this research does not lie in finished technology products, but in the success of translating empirical findings of user needs into targeted platform design requirements. Web-based HSK support is recommended to integrate the process of familiarizing test structures, specific skill training, formative feedback for follow-up, and self-monitoring features into a single complete digital ecosystem flow, so that the platform does not merely function as a question repository but as a structured training environment.

Several important limitations need to be noted in interpreting the results of this study. *First*, the sample size includes 50 respondents with a dominance of beginner teachers (with less than one year of service at 72%), so these findings are not intended to be widely generalized to the entire population of Mandarin language teachers or learners nationwide. *Second*, data were collected through a survey-based self-report method and analyzed descriptively, thus not testing the correlational relationship between variables nor predictive validity on official test scores. *Third*, the use of open-ended questions only serves to enrich the interpretation of quantitative findings, not to build a standalone qualitative model. *Fourth*, platform development at this stage is limited to formulating initial requirements for HSK Level 4 and has not yet been tested on completed prototype products. Therefore, the next research stages need to include content validation of assessments, *usability testing*, and empirical evaluation of the impact of platform usage on self-confidence, self-regulation, and user performance. With these limitations, this article remains within the corridor of needs analysis studies, while the design and evaluation of advanced products can be reported in separate publications.

ACKNOWLEDGMENTS

Thanks were extended to the Institute for Research and Community Service (LPPM) and the institution where the researcher is based for the moral support, facilities, and research permits granted. Sincere appreciation was also extended to all Mandarin language teachers and instructors in Indonesia who took valuable time to participate as respondents in filling out this needs analysis questionnaire. Constructive contributions and feedback from the participants became the main foundation in formulating the design specifications for the HSK 4 simulation platform in this study.

REFERENCES

Abdullaeva, B. S., Çakmak, F., & Abdullaev, D. (2024). Paper Assessment or Online Assessment: Exploring the Impact of Assessment Modes on EFL Students' Language Learning Outcomes and Personal Development. Language Testing in Asia, 14, 35. https://doi.org/10.1186/s40468-024-00309-w

Aljohani, N. J. (2025). Technology Mediated Task Based Language Teaching: A CALL Evaluation Framework and its Pedagogical Implications. Computer Assisted Language Learning, 38(3), 638–667. https://doi.org/10.1080/09588221.2024.2317840

https://pembahas.dialeks.id/index.php/jp
p-ISSN : 2809-5367
e-ISSN : 2809-6355

Bachman, L. F., & Palmer, A. S. (2010). Language Assessment in Practice: Developing Language Assessments and Justifying Their use in the Real World. Oxford University Press.

Brown, J. D. (2016). Introducing Needs Analysis and English for Specific Purposes. Routledge.

Chapelle, C. A. (2001). Computer Applications in Second Language Acquisition: Foundations for Teaching, Testing and Research. Cambridge University Press. https://doi.org/10.1017/CBO9781139524681

Chen, A., Zhang, Y., Jia, J., Liang, M., Cha, Y., & Lim, C. P. (2025). A Systematic Review and Meta-Analysis of AI-Enabled Assessment in Language Learning: Design, Implementation, and Effectiveness. Journal of Computer Assisted Learning, 41(1), e13064. https://doi.org/10.1111/jcal.13064

Creswell, J. W., & Creswell, J. D. (2023). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches (6th ed.). SAGE.

Estaji, M. (2024). Unveiling the Impact of Language Assessment Teacher Education on EFL Teachers' Assessment Agency and Professional Development Trajectory: Perceptions, Opportunities, and Challenges. Language Testing in Asia, 14, 20. https://doi.org/10.1186/s40468-024-00292-2

Fulcher, G. (2010). Practical Language Testing. Hodder Education.

Huang, L., & Zeng, J. (2025). Sustainability of Washback Effects: A Longitudinal Study of China's Twice-yearly NMET Reform. Language Testing in Asia, 15, 19. https://doi.org/10.1186/s40468-025-00358-9

Lee, S., & Jeon, J. (2024). Teacher Agency and ICT Affordances in Classroom-Based Language Assessment: The Return to Face-to-Face Classes After Online Teaching. System, 121, 103218. https://doi.org/10.1016/j.system.2023.103218

Li, Y., & Hu, B. (2025). Unpacking the Washback of the HSK on International Chinese Language Education: A Study of the Influencing Factors. Language Testing in Asia, 15, 31. https://doi.org/10.1186/s40468-025-00370-z

Ma, X. (2024). Enhancing Language Skills and Student Engagement: Investigating the Impact of Quizlet in teaching Chinese as a Foreign Language. Language Testing in Asia, 14, 5. https://doi.org/10.1186/s40468-024-00275-3

Maahs, I.-M., DeCapua, A., & Triulzi, M. (2025). Digital Technology and Language Learning: Insights from Teachers of Adult Migrant Learners. ReCALL, 37(2), 283–297. https://doi.org/10.1017/S0958344024000338

Morgana, V. (2024). Fostering English Speaking and Writing Subskills for the Cambridge B2 First Through Technology-Mediated Tasks. ReCALL, 36(2), 119–134. https://doi.org/10.1017/S0958344023000277

Sartika, S. B., Suyidno, & Wiguna, A. (2024). The Analysis of Students Needed in Digital Teaching Media. JTP – Jurnal Teknologi Pendidikan, 26(1), 44–62. https://doi.org/10.21009/jtp.v26i1.40737

Sawaki, Y., Ishii, Y., Yamada, H., & Tokunaga, T. (2024). Developing and Validating an Online Module for Formative Assessment of Summary Writing with Automated Content Feedback for EFL Academic Writing Instruction. Language Testing in Asia, 14, 50. https://doi.org/10.1186/s40468-024-00325-w

Shi, H., & Aryadoust, V. (2024). A Systematic Review of AI-Based Automated Written Feedback Research. ReCALL, 36(2), 187–209. https://doi.org/10.1017/S0958344023000265

Su, F., Zhao, Y., Zou, D., & Huang, B. (2026). WeChat-Based Collaborative Self-Regulated Learning Program: Exploring Self-Regulated Learning Strategy Use, Writing Performance, and Learning Behaviors. ReCALL. First View, 1–21. https://doi.org/10.1017/S0958344026100536

Ventivani, A., Sakti, K. F. L., Firdaus, D. O., Usadani, G. B., & Amri, M. (2026). The Impact of Digital Literacy on N§~„vuÍ‹ (Third-Level Vocabulary) in Mandarin Language Learners. JTP – Jurnal Teknologi Pendidikan, 28(1), 12–23. https://doi.org/10.21009/jtp.v28i1.57688

Vogt, K., Bøhn, H., & Tsagari, D. (2024). Language Assessment Literacy. Language Teaching, 57(3), 325–340. https://doi.org/10.1017/S0261444824000090

Wang, J., Zhou, H., Chen, S., Tong, H., & Yang, Y. (2024). How Teachers Support Secondary School Students to Become Self-Regulated Learners in Technology-Enhanced Language Learning. System, 123, 103313. https://doi.org/10.1016/j.system.2024.103313

Wang, J., Zheng, Y., & Zou, Y. (2024). Face Validity and Washback Effects of the Shortened PTE Academic: Insights from Teachers in Mainland China. Language Testing in Asia, 14, 32. https://doi.org/10.1186/s40468-024-00302-3

Xie, Q., & Jia, Q. (2025). Three Decades of Research on Washback (1993–2023): A Bibliometric Study. Language Testing in Asia, 15, 23. https://doi.org/10.1186/s40468-025-00357-w

Xu, Q., Sawaki, Y., & Zhu, Y. (2023). What Skills Are Being Assessed? Evaluating L2 Chinese Essays Written by Hand and on a Computer Keyboard. Assessing Writing, 57, 100765. https://doi.org/10.1016/j.asw.2023.100765

Yang, C. T.-Y., & Chen, H. H.-J. (2026). ChatGPT and L2 Chinese writing: Evaluating the Impact of Model Version and Prompt Language on Automated Corrective Feedback. Computer Assisted Language Learning, 39(4), 1119–1147. https://doi.org/10.1080/09588221.2025.2453205

Yang, Y., & Jirawit, Y. (2025). Navigating HSK level 5 Challenges: Dual Prspectives from Thai Undergraduates and Chinese Lecturers. Language Testing in Asia, 15, 24. https://doi.org/10.1186/s40468-025-00360-1

Ye, X., Yang, Y., Qie, Y., & Hu, Z. (2024). The Impact of 24-h Take-Home Exam on Language Learning and Teaching on the China Campus of a British University. Language Testing in Asia, 14, 23. https://doi.org/10.1186/s40468-024-00295-z

Yu, C., & Xu, W. (2024). Writing Assessment Literacy and Its Impact on the Learning of Writing: A Netnography Focusing on Duolingo English Test Examinees. Language Testing in Asia, 14, 24. https://doi.org/10.1186/s40468-024-00297-x

Zhang, E. D., & Yu, S. (2024). Understanding L2 Student Writers' Self-Assessment in Digital Multimodal Composing: A Process-Oriented Approach. System, 121, 103219. https://doi.org/10.1016/j.system.2024.103219

Zhang, W., Liu, B., & Wilson, A. J. (2024). Examining Chinese EFL Learners' Online Self-Regulated Learning: A Mixed-Methods Approach. System, 123, 103277. https://doi.org/10.1016/j.system.2024.103277

Zhang, Y., Yang, F., Yang, H., & Han, S. (2024). Does Checking-In Help? Understanding L2 Learners' Autonomous Check-In Behavior in an English-Language MOOC Through Learning A nalytics. ReCALL, 36(3), 343–358. https://doi.org/10.1017/S0958344024000144