

System for Competency Assessment in Language Education (SCALE): A Computer-Mediated Assessment System for Measuring Proficiency in Indian Languages

Pankaj Dwivedi ¹, , Amit Kumar Jha ²

1. National Testing Service, Central Institute of Indian Languages, Mysuru, IND

2. National Testing Service-India, Central Institute of Indian Languages, Mysuru, IND

Received: September 13, 2025 | Review began: September 23, 2025 | Review ended: April 13, 2026 | Published: April 23, 2026

© Copyright 2026

This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

This paper explains the development, including design, implementation, and administration, of a semi-automated web application named *System for Competency Assessment in Language Education* (SCALE) for assessing proficiency in Listening, Speaking, Reading, and Writing (LSRW) skills in Indian languages. The database utilised for the task comprised 3,09,252 question items stored in text, image, and audio formats and marked for A1 to C2 levels to test specific language skills such as pronunciation, spelling, grammar, meaning, and comprehension at various difficulty levels of the concerned languages. The test items have been created by experts in 22 languages using materials such as prescribed textbooks (453), various national and state-level newspapers and magazines, while ensuring that the culture-specific nuances remain intact. This application is robust, scalable, and secure. It can be accessed by multiple users simultaneously without any physical or temporal limitations. The user evaluation of the application stands to the mean assessment score of the LSRW skills of 74.7%. In comparison, 80.4% of the participants express that the given questions are duly inclusive of the respective culture associated with the language. The importance of the application lies in the fact that this is the first web application designed to assess all four linguistic skills of proficiency (LSRW) in all 22 scheduled languages of India. The web application can be expanded to any number of languages.

Categories: Computational Linguistics , Digital Humanities, Web applications

Keywords: automated educational assessment, indian languages, language proficiency, language testing, lsrw skills, online test

Introduction

India is a multilingual and multicultural country with a population exceeding 1.4 billion. National Education Policy - 2020 of India puts unprecedented emphasis on the teaching of and teaching in Indian languages at all identified levels of education. More and more central and state organizations and agencies have started conducting entrance tests and recruitment examinations in major languages of India such as National Testing Agency, Nationalized Banks, and Staff Selection Commission. The initiative has been taken into account to provide the stakeholders with equal opportunity in admission and employment while reducing linguistic discrimination and dependence on English. Whereas these initiatives have primarily set the correct path for the promotion, recognition, and development of Indian languages, they also have revealed that India does not yet have any central or accepted mechanism, framework, or tool for the valid assessment of the proficiency of the candidates at any level. Due to a lack of expertise and resources, admission and recruitment agencies often provide their own assessment tests to the candidates, which frequently are invalid, extra- or under-focused on a particular language skill (e.g., reading or writing), and skewed to the personal judgment of test

How to cite this article:

Dwivedi P, Jha A (April 23, 2026) System for Competency Assessment in Language Education (SCALE): A Computer-Mediated Assessment System for Measuring Proficiency in Indian Languages. *Cureus J Comput Sci* 3 : es44389-025-00055-x. DOI <https://doi.org/10.7759/s44389-025-00055-x>

makers. These tests, therefore, do not lead to a fair evaluation of test takers for their linguistic skills, and they further compromise the overall evaluation process of an assessment test. Additionally, administering these traditional tests is challenging due to the test takers' temporal, geographical, and financial limitations. Kumar [1] states similar reasons for the popularity of Open and Distance Learning and online programs in India.

Due to practical reasons, there have been many attempts to develop computerized assessment systems for assessment of each language skill, i.e., listening comprehension [2,3], oral and spoken proficiency [4,5], reading [6,7], and writing proficiency [8,9], using the popular natural language processing and computational methods. Chalhoub-Deville [10] observes that the computerized delivery of tests is appealing, effective, and viable for the administration of standardized language tests. A developing country like India, where, on the one hand, states and union territories may have multiple languages as the medium of instruction or as subjects of education, severely lacks the financial and human resources needed to cater to the teaching and learning needs of learners. On the other hand, online proficiency assessments are often preferred over face-to-face ones for self-explanatory reasons.

Language proficiency frameworks

Most admission and recruitment organizations and agencies describe language proficiency on a three-point scale, Basic, Intermediate, and Advanced. Though these are often vaguely defined terms for the correct demarcation of a person's ability to use a language in professional settings, they have already secured their place not only in public discourse but also in industry and academia. Traditionally, teachers, test makers, and evaluators frequently use these terms while assessing a person's ability to understand and use a language for various purposes. There are a few recognized frameworks, such as the Common European Framework of Reference for Languages (CEFR) [11], the American Council on the Teaching of Foreign Languages Proficiency Guidelines (ACTFL) [12], Canadian Language Benchmarks (CLB) [13], and Interagency Language Roundtable scale (ILR) [14], for the assessment of proficiency in the English language. Since English is very close in terms of genesis and structure, these frameworks sometimes are also applied in the assessment of other European languages such as French, German, Italian, Polish, and Spanish. Of the above frameworks, CEFR has earned recognition and broader acceptance on international platforms. The CEFR claims to be not tied to any specific language test, and most standardized language tests now provide CEFR-level equivalents. The CEFR describes language ability on a six-point scale (A1, A2, B1, B2, C1, and C2), from A1 for beginners up to C2 for those who have mastered a language. Other than Japanese Language Proficiency Test by Japan Foundation and Japan Educational Exchanges and Services (JEES), TOPIK (Test of Proficiency in Korean) National Institute for International Education (NIIE), under South Korea's Ministry of Education. Chinese Proficiency Test (HSK) by CTI Co. Ltd., etc., is used by the respective governments of Japan, Korea, and China. Table 1 shows the comparative mapping of proficiency levels across the three frameworks: General, CEFR, and National Testing Service-India (NTS-I) frameworks.

How to cite this article:

Dwivedi P, Jha A (April 23, 2026) System for Competency Assessment in Language Education (SCALE): A Computer-Mediated Assessment System for Measuring Proficiency in Indian Languages. *Cureus J Comput Sci* 3 : es44389-025-00055-x. DOI <https://doi.org/10.7759/s44389-025-00055-x>

General	CEFR	NTS-I
Basic	A1	Primary
	A2	Upper Primary
Intermediate	B1	Secondary
	B2	Senior Secondary
Advanced	C1	Graduation
	C2	Post-Graduation
		Research

TABLE 1: Comparative mapping of the various proficiency levels across the three frameworks

CEFR, Common European Framework of Reference for Languages; NTS-I, National Testing Service-India

As for Indian languages, NTS-I of the Central Institute of Indian Languages has developed a framework, known as System for Competency Assessment in Language Education (SCALE), which is graded into seven levels of different proficiency: Primary, Upper Primary, Secondary, Higher Secondary, Graduation, Post-Graduation, and Research. Whereas Primary level refers to the first stage, or the beginner, research level refers to proficiency when a test taker is able to communicate complex ideas after mastering the language [15].

Background: Use of technology in language testing

Technology has influenced all disciplines of inquiry, including language education. In the past two decades, technology has phenomenally impacted teaching-learning, testing, and assessment of languages. The concerted efforts on the integration of technology into the field of language testing can be traced back to the Language Testing Research Colloquium organized at Educational Testing Service in Princeton in 1985 [16]. The colloquium saw many papers dealing with computer-based testing and computer adaptive testing, with their latent trait models focusing on item-bank construction, item selection, and scoring procedures. Another example was a program available on British Council computers that enabled test developers and classroom teachers to create gap-filling tests [17]. This was considered cutting-edge computer technology at the time and one of the first applications of item response theory in language testing for estimating test-takers' levels of ability. Following that, the field has seen tremendous development, especially after the recent advances in the area of machine learning and artificial intelligence.

There have been many platforms, web applications, and software that offer specialized language learning courses in many languages of the world via online mode, but few offer testing and assessment services for learners. The most notable organization in this area is the Educational Testing Service, USA. And, the most popular tests are the International English Language Testing System (IELTS) and the Test of English as a Foreign Language (TOEFL) conducted to assess the English proficiency of foreign nationals to be eligible to study, work, or reside in English-speaking countries such as USA, UK, Australia, Canada, and New Zealand. European Union via the EU academy offers assessment tests in many other languages of Europe such as Bulgarian, Hungarian, Greek, Estonian, Latvian, Portuguese, Slovak, Romanian, and Croatian [18] to help people understand their present level of proficiency in the concerned language. Similarly, the Chinese Proficiency Test (HSK), administered by CTI Co. Ltd. (CTI), is an international standardized test for Chinese

How to cite this article:

Dwivedi P, Jha A (April 23, 2026) System for Competency Assessment in Language Education (SCALE): A Computer-Mediated Assessment System for Measuring Proficiency in Indian Languages. *Cureus J Comput Sci* 3 : es44389-025-00055-x. DOI <https://doi.org/10.7759/s44389-025-00055-x>

language proficiency [19]. European countries such as France (TEF IRN), Germany (Goethe-Zertifikat, A1-C2), Poland (państwowe egzaminy certyfikacyjne z języka polskiego jako obcego), and Spain (Diplomas de Español como Lengua Extranjera) also conduct their own language tests for assessing the suitability of foreign nationals willing to reside there for various purposes [20-22]. These tests or certifications usually are based on the CEFR published by the Council of Europe. Along with the assessment of the Listening, Speaking, Reading, and Writing (LSRW) skills, these tests also aim to assess understanding of the culture associated with the concerned linguistic community.

SCALE

For ease of convenience and acceptability, in SCALE portal, we equated General Framework, CEFR, and NTS-I frameworks in a single window system so that a wide range of stakeholders can make use of the portal to get a valid assessment of proficiency in any of the Indian languages. There is a provision of selection from the given frameworks and corresponding levels within according to the user's choice and requirement. Based on their framework and level selection, the algorithm randomly generates the test items without any repetition. For randomized selection, a C#-based algorithm is employed that mimics the functionality of a RAND function. In practice, the algorithm utilizes a pseudo-random number generator (Random class in C#) to assign random indices to the question bank and selects the required number of questions without repetition. This ensures that each question item has an equal probability of being chosen, maintaining test fairness and validity. The typical implementation uses an approach similar to reservoir sampling or random indexing:

```
var rng = new Random();  
var sample = questions.OrderBy(q => rng.Next()).Take(n).ToList();
```

This ensures that from the large dataset, a unique, reproducible, and statistically fair subset of questions is generated for each test instance. From an assessment perspective, such randomization is critical because it reduces predictability, prevents item exposure, and ensures that no two candidates receive identical test forms. By accurately reflecting the candidate's proficiency across diverse items, this not only upholds test validity but also enhances reliability, since the random selection balances the distribution of items across skill areas and difficulty levels. There have been a few studies on the automatic generation of question papers and test items in various academic fields [23-25]. However, most of these question papers focus on generating objective type of question papers and are often monolingual. The design and development of a multilingual database meant for the assessment of proficiency of LSRW skills are more complex in nature, not only in terms of the system architecture but also in the creation of the question database.

The SCALE platform presently assesses four core language skills: LSRW. For A1 and A2 levels, a complete test takes approximately 100 minutes, with 30 minutes each allotted for Listening and Writing, and 20 minutes each for Speaking and Reading. These time allocations, however, may be modified as per the requirements of the respective agencies using the system for assessment of different levels of proficiency. The platform is web-based and can be accessed from any computer with a modern browser and a reliable internet connection. For the Listening and Speaking components, it is essential to have a good-quality headset equipped with a functional microphone. At present, SCALE does not offer Remote Proctoring Assessment owing to security concerns. Therefore, it is recommended that the tests be conducted in a controlled environment under the supervision of a human invigilator to ensure fairness, integrity, and compliance with testing standards.

Table of specification

First, to maintain the quality of the academic test items, it was decided to collect language textbooks prescribed up to the secondary level of various state and central government boards. A total of 453 prescribed textbooks, various newspapers and magazines were collected for 22 scheduled languages. The contents of the books were further studied to understand the form, structure, vocabulary, grammatical, and cultural aspects of the concerned languages. Wherever needed, more text was sourced from the newspapers, websites, and other books. The language experts then created a common template with varying linguistic idiosyncrasies of the respective languages. Accordingly, a content-free sample format of the question paper across the most frequently asked questions containing 100 questions were prepared. This

How to cite this article:

Dwivedi P, Jha A (April 23, 2026) System for Competency Assessment in Language Education (SCALE): A Computer-Mediated Assessment System for Measuring Proficiency in Indian Languages. *Cureus J Comput Sci* 3 : es44389-025-00055-x. DOI <https://doi.org/10.7759/s44389-025-00055-x>

sample format was common to all 22 languages. Since listening and reading are more primary skills than speaking and writing, respectively, it was agreed to give more weightage to them in terms of the number of questions asked to assess these skills. Also, the least weightage was given to the writing skills to minimise any significant deviation arising due to the type of keyboard used by the user as assessment of writing skills in any computer-mediated assessment inherently refers to type-writing assessment in principle. The same timeframe is allotted to assess the two complementary skills, i.e., listening & speaking, and reading & writing. Table 2 provides the overview of the content-free specifications of the test with 100 questions meant to assess LSRW. It also gives a clear description of the type of question items and tasks required to be completed by the test takers.

	Listening	Speaking	Reading	Writing
	Select the word you hear	Speak the given word	Spelling correction	Write about a given topic
Number of Questions	8	8	8	2
	Select the sentence you hear	Speak the statement	Sentence correction	Summarise the paragraph in 50 words
Number of Questions	4	8	4	2
	Complete the sentence	Speak about a topic	Poem	Arrange the sentence in the correct order
Number of Questions	4	2	4	8
	Short conversation	Describe the picture	Story	Sentence transformation
Number of Questions	6	2	6	4
	Listen to the passage and answer the questions	Introduce yourself	Comprehension passage	Describe the picture
Number of Questions	8	1	8	1
		Talk about the language		Write about yourself/family member
		1		1
Total	30	22	30	18

TABLE 2: Overview of the content-free specifications of the test with 100 questions meant to assess Listening, Speaking, Reading and Writing

How to cite this article:

Materials And Methods

System management

The architecture of the system is divided into four integral modules: User Management, Test Management, Question Bank Management, and Result Management. Each of these modules addresses a distinct yet complementary dimension of the system's functionality. The User Management component ensures secure access by authenticating user credentials while simultaneously enabling the registration of new users. It, thereby, helps maintain both integrity and scalability of the platform. The Test Management module encompasses the creation, scheduling, and regulation of tests, which allows administrators to define assessment parameters, revise, or remove outdated tests, and facilitate the actual conduct of examinations. The Question Bank Management subsystem constitutes the intellectual repository of the system. Here, questions can be added, edited, or deleted in order to preserve academic rigor and adaptability. The Result Management module records the outcomes of assessments, enables their retrieval, and performs analytical functions to extract meaningful pedagogical insights. These interdependent modules together create an integrated and reliable framework. In sync with one another, they ensure not only operational efficiency but also a continuous feedback loop for enhancing learning outcomes. Figure 7 depicts the integrated flow of four modules, i.e., User Management, Test Management, Question Bank Management, and Result Management, working in sync with one another to guide the stakeholders through the entire examination process.

How to cite this article:

Dwivedi P, Jha A (April 23, 2026) System for Competency Assessment in Language Education (SCALE): A Computer-Mediated Assessment System for Measuring Proficiency in Indian Languages. *Cureus J Comput Sci* 3 : es44389-025-00055-x. DOI <https://doi.org/10.7759/s44389-025-00055-x>

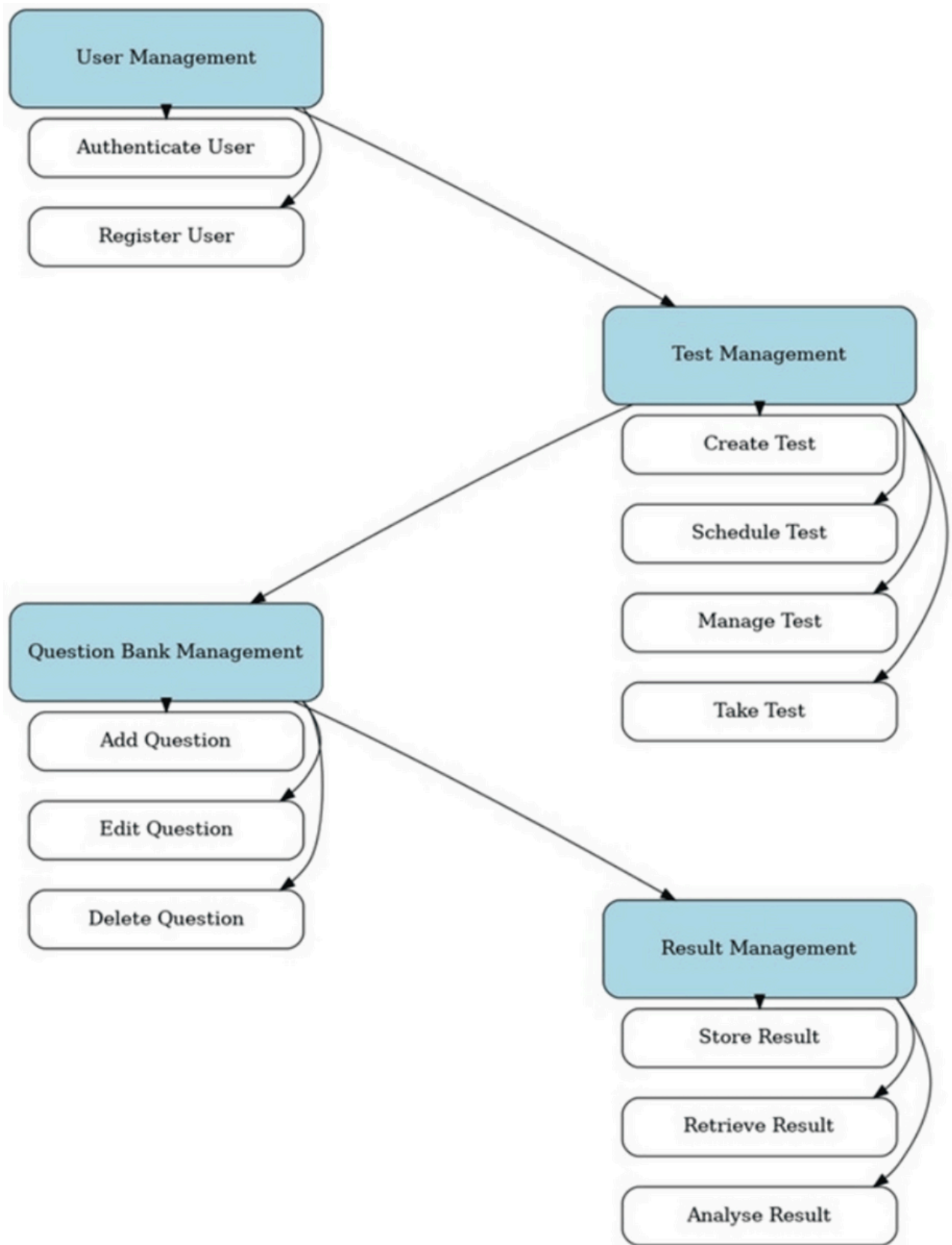


FIGURE 1: System management

How to cite this article:

Dwivedi P, Jha A (April 23, 2026) System for Competency Assessment in Language Education (SCALE): A Computer-Mediated Assessment System for Measuring Proficiency in Indian Languages. *Cureus J Comput Sci* 3 : es44389-025-00055-x. DOI <https://doi.org/10.7759/s44389-025-00055-x>

System architecture: inputting, retrieval, and assessment module

The system architecture consists of three modules, namely, the Input Module, Retrieval Module, and Assessment Module, with their respective front and backend applications. Microsoft SQL server is used to design the database, i.e., tables and stored procedures. Active Server Page Extended (ASPX) framework with C# and Bootstrap 5 has been used to design typography, navigation, dashboard, forms, and other interface components to connect them with varying sizes of screens to enable users to attempt assessment tests from different devices such as computers, mobile phones, and tablets. The database comprises several tables used for managing multilingual assessment data (Table 3).

Table Name	Description
Domain	To maintain different domains and their ID for language test
Media	To maintain different media files such as images and audio
QuestionTable	To store questions and other details related to questions
UT	To store user details
TTT	To maintain questions and other details related to the test, such as questions, session ID, date of test, etc.
Tracking User Details	To maintain the record of results and marks of different skills of the examinee

TABLE 3: Names and descriptions of the various tables used in the preparation of the multilingual database

Inputting and Retrieval Modules

Inputting and retrieval modules consist of three screens: the Media Screen, Question Screen, and View/Edit screen. The media screen lets the writer upload the required image, audio, or video file to the question items in valid formats such as mp3, jpg/jpeg, and mp4. Then, the question item is written by an item writer following the table of specifications specifying the dimensions such as domain, type, subcomponents, instructions, body, valid options, answer key or cue, marks, and maximum time. Inputted question is then stored in the database. A user may log in through credentials and perform the assigned tasks based on the different roles such as item writer, question item validator, examinee, and evaluator. The inputted question items may be viewed and edited as per requirement by an assigned validator (Figure 2 and Figure 3). In the end, a web-based SAP crystal application is used for the purpose of generating a final report to inform the users about the assessment performance. Figure 4 and Figure 5 present sequential inputting of test question items as they appear on the front end to the users in their assigned roles.

How to cite this article:

Language Proficiency Test Platform
Hi, Ms.Lega.S!

View Question Items

Click to View All Question Items

BACK

Total Question Item: 1142

Question Item Details

Show 10 entries Search:

Sl. No.	SubComponent	Questions	Options1	Options2	Options3	Options4	Answer	Edit QI	View QI	Va
275	1100-Select the word you hear	நீங்கள் கேட்கும் வார்த்தையை தேர்வு செய்க	பட்டு	விட்டு	தட்டு	சொட்டு	தட்டு	Edit	View	0
276	1100-Select the word you hear	நீங்கள் கேட்கும் வார்த்தையை தேர்வு செய்க	சிங்கம்	தங்கம்	புங்கம்	சுங்கம்	தங்கம்	Edit	View	0
277	1100-Select the word you hear	நீங்கள் கேட்கும் வார்த்தையை தேர்வு செய்க	நாக்கு	பாக்கு	மூக்கு	வாக்கு	மூக்கு	Edit	View	0
278	1100-Select the word you hear	நீங்கள் கேட்கும் வார்த்தையை தேர்வு செய்க	பட்டை	சட்டை	மட்டை	தட்டை	சட்டை	Edit	View	0
279	1100-Select the word you hear	நீங்கள் கேட்கும்	அருவி	குருவி	மருவி	உருவி	அருவி	Edit	View	0

FIGURE 2: Question view screen

How to cite this article:

Language Proficiency Test Platform Hi, Ms.Archita Gogoi!

Question Item Edit Screen

BACK

Edit Question Screen. Please edit required fields.

Domain: General

MediaType: Passage

MediaTitle: RStory_XahokhiGhura

Image/Audio: Click to select Image/Audio File:

Passage/Poem:

Edit Format Table

Paragraph B I U [Text alignment icons]

সাহসী ঘোঁৰা

১৬ শতিকাৰত ৰাজস্থানত ৰাণা প্ৰতাপ নামৰ এজন মহান যোদ্ধা ৰজা বাস কৰিছিল। তেওঁ মোৰাৰত শাসন কৰিছিল, যি সাহস আৰু বীৰত্বৰ বাবে পৰিচিত ৰাজ্য আছিল। ৰাণা প্ৰতাপ এজন শক্তিশালী আৰু নিৰ্ভীক ৰজা আছিল যদিও তেওঁ বহু প্ৰত্যাহ্বানৰ সন্মুখীন হৈছিল, বিশেষকৈ তেওঁৰ ভূমি দখল কৰিবলৈ বিচৰা মোগল সাম্ৰাজ্যৰ পৰা।

১৫৭৬ চনত হালদীঘাটত ৰাণা প্ৰতাপৰ বাহিনী আৰু মোগল সেনাৰ মাজত এক বিশাল যুদ্ধ সংঘটিত হয়। সংখ্যা বেছি হোৱাৰ পিছতো ৰাণা প্ৰতাপে সাহসেৰে যুঁজিছিল, সন্মুখৰ পৰা নিজৰ সৈন্যক নেতৃত্ব দিছিল। গোটেই যুদ্ধত তেওঁৰ নিষ্ঠাবান ঘোঁৰা চেতক তেওঁৰ লগত আছিল। চেতক গতি, শক্তি আৰু আনুগত্যৰ বাবে পৰিচিত আছিল।

যুদ্ধৰ সময়ত ৰাণা প্ৰতাপ আহত হৈছিল আৰু বন্দী হোৱাৰ আশংকাত ভুগিছিল। মোগল সৈন্যে যেতিয়া তেওঁৰ ওচৰ চাপিছিল লগে লগে চেতকে এখন নৈৰ ওপৰেৰে বীৰত্বপূৰ্ণ জাঁপ মাৰি ৰাণা প্ৰতাপক নিৰাপদ স্থানলৈ লৈ যায়। কিন্তু তেনে কৰাৰ ফলত চেতকও গুৰুতৰভাৱে আঘাতপ্ৰাপ্ত হয়। মালিকক সুৰক্ষিত হোৱাটো নিশ্চিত কৰাৰ পিছত চেতকে শেষ নিশ্বাস এৰি চলি পৰিল।

SubComponent: 3400-Story

Education Level: Upper Primary Item Format: Multiple Choice

Question Instruction: নিম্ন গল্পটো পঢ়ি তলৰ প্ৰশ্নৰ উত্তৰ দিয়ক

Question Item:

Edit Format Table

Paragraph B I U [Text alignment icons]

হালদীঘাট যুদ্ধৰ সময়ত কি ঘটিছিল যিয়ে চেতকক কিংবদন্তি কবি তুলিছিল?

Option1: চেতক আহত যদিও কোনোমতে পলায়ন কৰি Option2: চেতকে সেনাবাহিনীক বিজয়ৰ দিশে আগবঢ়া

Option3: ৰাণা প্ৰতাপক বচাবলৈ চেতকে বীৰত্বপূৰ্ণ জাঁপ Option4: চেতকে মোগল সৈন্যৰ লগত যুদ্ধ কৰিছিল

Answer: ৰাণা প্ৰতাপক বচাবলৈ চেতকে বীৰত্বপূৰ্ণ জাঁপ মাৰিছিল

Marks: 1 Time: 1

Remarks (During Edit): Mention the remarks for this question if required

Update Question Item

How to cite this article:

FIGURE 3: Question edit screen

Language Proficiency Test Platform

Hi, Ms.Archita Gogoi !

Media Screen. Please input fields to create media.

Select Domain:

Media Type:

Title:

Stay on Social Media

- Facebook
- Twitter
- LinkedIn
- Youtube

Quick Links

- > Login
- > SignUp
- > About Us
- > Faq

Address

National Testing Service-India
Center for Testing & Evaluation,
Central Institute of Indian Languages,
Manasagangotri, Mysuru-570006

Contact Us

0821-2515006
headnts@gmail.com
reach.nts.india@gmail.com

Copyright © 2024, National Testing Service-India, CIL, Mysuru

Designed by NTS-I

FIGURE 4: Media inputting screen of SCALE

How to cite this article:

Dwivedi P, Jha A (April 23, 2026) System for Competency Assessment in Language Education (SCALE): A Computer-Mediated Assessment System for Measuring Proficiency in Indian Languages. *Cureus J Comput Sci* 3 : es44389-025-00055-x. DOI <https://doi.org/10.7759/s44389-025-00055-x>

Language Proficiency Test Platform

Hi, Ms.Archita Gogol !

Question Input Screen. Please input fields to form question.

Select Domain:

Select Title:

Select Media: Select Passage:

Level of Proficiency:

Select ItemFormat:

Select Question Type:

Question Instruction:

Question Item:

Edit Format Table
↶ ↷ **B** *I* U

Options:

Answer / Answer Cues:

Marks:

Time in mins:

Stay on Social Media

- Facebook
- Twitter
- LinkedIn
- Youtube

Quick Links

- Login
- SignUp
- About Us
- Faq

Address

National Testing Service-India
Center for Testing & Evaluation,
Central Institute of Indian Languages,
Manasagangotri, Mysuru-570006

Contact Us

- 0821-2515006
- headnts@gmail.com
- reach.nts.india@gmail.com

FIGURE 5: Question inputting screen

How to cite this article:

Dwivedi P, Jha A (April 23, 2026) System for Competency Assessment in Language Education (SCALE): A Computer-Mediated Assessment System for Measuring Proficiency in Indian Languages. Cureus J Comput Sci 3 : es44389-025-00055-x. DOI https://doi.org/10.7759/s44389-025-00055-x

Question item writers add new questions using the question inputting system and after inputting the questions if (s)he finds any errors (s)he can edit those questions. Question item writers do not have permission to delete a question. Only the administrator can delete a question based on the requirement. As this system is semi-automated, objective-type questions are evaluated by the system itself, while the system routes subjective assessment items and stored responses to the assigned human evaluators. Human evaluators are provided with a set of response rubrics and answer cues to evaluate the responses against the subjective type of questions for the assigned marks.

Evaluation Module

The evaluation module consists of four time-locked assessment tests to assess LSRW skills. These four tests are arranged in order of LSRW as per the standard norms of the field. However, a test taker may choose any of the tests depending on her/his comfort. However, once chosen, the user cannot switch over to another test without submitting the previous one. The users will need common accessories such as a standard speaker to attempt a listening assessment and a microphone to attempt a speaking assessment. Also, the users must have unicode-compliant fonts of the concerned language and compatible keyboards installed in their respective systems. After the successful submission of all assessment tests, the users immediately receive partial scores against the objective type of question items, while answers submitted against a subjective type of question item are safely routed to the assigned evaluator. The human evaluator based on the rationality, validity, and correctness of the answer assigns marks to a user. In the end, the final score is calculated by the system and sent back to the user.

To maintain consistency during the assessment process, human evaluators are directed to follow a structured set of scoring criteria supported by detailed rubrics and answer cues. While this helps bring a degree of uniformity during the evaluation process, the involvement of human evaluators may still introduce variation and limit scalability in larger deployments. It may also be noted that fully automated assessment of subjective responses, particularly for speaking and writing, remains limited in the context of Indian languages due to resource constraints involved. The steps such as evaluator training and moderation are being considered to further strengthen the process.

System algorithm and flowchart

The algorithm for the semi-automated test system has been designed to ensure a structured and pedagogically sound assessment process. It begins with the initialization of variables, data structures, and relevant information such as test details, questions, and user profiles. The procedure commences with user authentication, wherein individuals are prompted to log in, and their credentials are validated, with provisions for retry or registration in cases of failure. Once authenticated, the system guides the user to select the language of assessment from a predefined list, followed by the choice of an appropriate proficiency framework such as CEFR, NTS-I, or a general scale. Within each framework, proficiency levels (ranging from beginner to advanced, or from A1 to C2, as well as corresponding academic stages) and the specific linguistic skills - LSRW - are identified for testing. Based on these inputs, the system configures the assessment by retrieving suitable test items from the database and determining the duration according to the skill selected. The system, presently, draws from a database of over 3,09,252 test items across 22 languages, including text, audio, and image-based content. The retrieval of test items is carried out in real time, and no noticeable delay is experienced during test generation.

The assessment phase begins with the presentation of instructions and initiation of the test timer, after which items are displayed sequentially and randomly. It is ensured that no repetition happens within the same task. Questions are delivered in appropriate formats, allowing users to input their responses, which are then submitted skill by skill until the completion of the assessment. Upon final submission or expiration of time, the system halts the timer, prompts for confirmation, and initiates the evaluation process. For objective items, the system automatically assigns scores based on correctness, while subjective responses are routed to designated evaluators with explicit assessment guidelines. Once evaluation is complete, the system presents the results by displaying scores and proficiency levels in the selected language. Finally, users are offered options to review performance, access certificates (if applicable), log out, or explore other sections of the system. This structured algorithm thus integrates authentication, configuration, assessment, evaluation, and results into a coherent framework for conducting language proficiency testing (Figure 6).

How to cite this article:

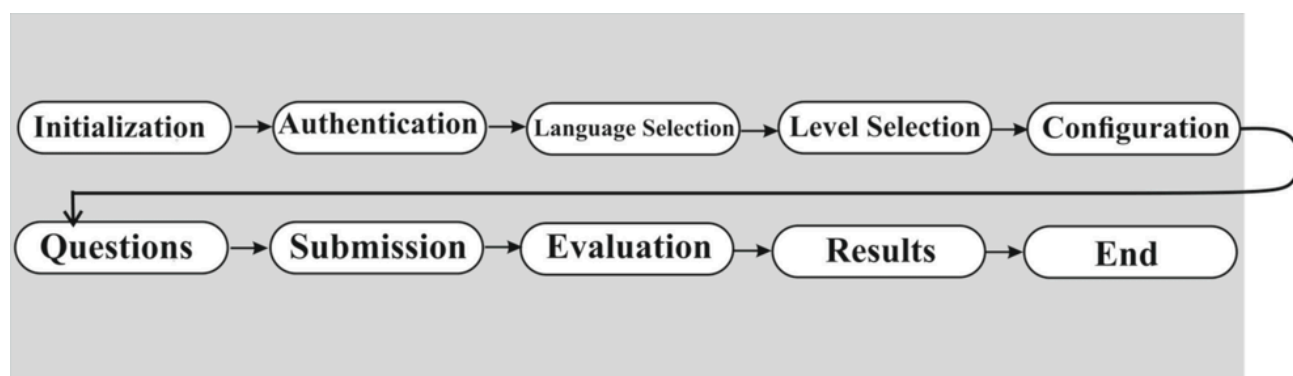


FIGURE 6: System flowchart

This algorithm outlines the process of conducting a language proficiency test, from the authentication of the users to the presentation of the result. It includes steps for selecting the language and proficiency level, configuring the test, presenting questions, evaluating answers, and providing feedback.

Results And Discussion

User evaluation

The application was subjected to the overall evaluation by 110 users in total, representing 5 users for each language. The users were asked to assess the skill wise, i.e., listening, speaking, reading and writing, performance of the application, and its accurate representation of the associated culture. The users involved the students, staff, and faculty members of different Schools/Institute/Colleges who participated in a week-long three training programs focusing on the testing and evaluation of Indian languages at NTS-I, Central Institute of Indian Languages. They also differed in terms of their exposure to computer-based testing.

The participants were asked to respond using a simple three-point scale (Yes/No/Somewhat) across five questions on their overall experience of the assessment application (Table 4):

Question 1 - Do you think the present test correctly measures your Listening proficiency in the given language?

Question 2 - Do you think the present test correctly measures your Speaking proficiency in the given language?

Question 3 - Do you think the present test correctly measures your Reading proficiency in the given language?

Question 4 - Do you think the present test correctly measures your Writing proficiency in the given language?

Question 5 - Do you think the application correctly represents the cultural aspects associated with the given language?

How to cite this article:

Dwivedi P, Jha A (April 23, 2026) System for Competency Assessment in Language Education (SCALE): A Computer-Mediated Assessment System for Measuring Proficiency in Indian Languages. *Cureus J Comput Sci* 3 : es44389-025-00055-x. DOI <https://doi.org/10.7759/s44389-025-00055-x>

Skill	Response	Percentage	Skill	Response	Percentage
1 - Listening			3 - Reading		
Yes	61	55.5	Yes	102	92.7
No	29	26.3	No	3	2.7
Somewhat	20	18.2	Somewhat	5	4.5
2 - Speaking			4 - Writing		
Yes	67	60.9	Yes	99	90
No	26	23.6	No	5	4.5
Somewhat	17	15.5	Somewhat	6	5.5
LSRW – mean score	Response	Percentage	5 - Cultural	Response	Percentage
Yes	329	74.7	Yes	354	80.4
No	63	14.3	No	17	3.8
Somewhat	48	10.9	Somewhat	69	15.6

TABLE 4: Skill-wise user evaluation

The three-point scale format was chosen to keep the task simple and straightforward, especially, since the participants came from different linguistic and cultural backgrounds and had varying levels of familiarity with such online systems.

It is observed that the overall mean satisfaction of the assessment score of the LSRW skills stands to be 74.7%, while 80.4% of the participants expressed that given questions are duly inclusive of the respective culture associated with the language. It is also noted that listening and speaking skills have scored the minimum percentage, i.e., 55.4% and 60.9%, respectively. This aspect requires further investigation through more detailed and controlled studies. Lowest scores on these two skills are due to the fact that users, in most cases, are more comfortable with the computer-mediated reading and writing tests than listening and speaking.

Conclusions

The application architecture has been designed in such a way that new question items across various dimensions can further be modified to cater to the specific needs, i.e., be it the inclusion of a new language, addressing an inadvertent error in an existing question item, or creation of new test item to check specific linguistic skills of test takers. Also, the system can administer assessment tests as per the defined needs of the users, ranging from the number of question items per assessment set to the allotted time, freedom to select the order of skills across the four levels, fully automated evaluation of the objective type of the question items presenting immediate scores, while safe routing the recorded

How to cite this article:

responses of the subjective type of test items to the assigned human evaluators. This online semi-automated application is robust, scalable, and secure in nature. It can be accessed by users at their convenience. It also issues an e-certificate of the scores upon final evaluation of the assessment items. It is recommended that the system be used for recruitment and admissions purposes in controlled environments, subject to further large-scale validation and standardisation.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Pankaj Dwivedi, Amit Kumar Jha

Acquisition, analysis, or interpretation of data: Pankaj Dwivedi, Amit Kumar Jha

Drafting of the manuscript: Pankaj Dwivedi, Amit Kumar Jha

Critical review of the manuscript for important intellectual content: Pankaj Dwivedi, Amit Kumar Jha

Supervision: Pankaj Dwivedi

Disclosures

Human subjects: All authors have confirmed that this study did not involve human participants or tissue. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Acknowledgements

The authors acknowledge the academic and technical support provided by various language and technical experts and resource persons of the National Testing Service - India and the Central Institute of Indian Languages.

References

1. Kumar MJ: [The emergence of open and distance learning \(ODL\) and online programmes in India](#). IETE Technical Review. 2024, 41:121-122. [10.1080/02564602.2024.2326750](https://doi.org/10.1080/02564602.2024.2326750)
2. de Wet F, Van der Walt C, Niesler TR: [Automatic assessment of oral language proficiency and listening comprehension](#). Speech Communication. 2009, 51:864-874. [10.1016/j.specom.2009.03.002](https://doi.org/10.1016/j.specom.2009.03.002)
3. de Wet F, Müller PDV, Van der Walt C, Niesler TR: [Experiments in automatic assessment of oral proficiency and listening comprehension for bilingual South African speakers of English](#). Proceedings of the Nineteenth Annual Symposium of the Pattern Recognition Association of South Africa. 2008, 67-72.
4. Bernstein J, De Jong J, Pisoni D, Townshend B: [Two experiments on automatic scoring of spoken language proficiency](#). STIL2000 (Integrating Speech Technology in Learning). International Speech Communication Association, Dundee, UK; 2000. 1:57-61.
5. Kang O, Johnson D: [The roles of suprasegmental features in predicting English oral proficiency with an automated system](#). Language Assessment Quarterly. 2018, 15:150-168. [10.1080/15434303.2018.1451531](https://doi.org/10.1080/15434303.2018.1451531)
6. Gerosa M, Narayanan S: [Investigating automatic assessment of reading comprehension in young children](#). IEEE International Conference on Acoustics, Speech and Signal Processing, Las Vegas, NV, USA. 2008, 5057-5060. [10.1109/ICASSP.2008.4518795](https://doi.org/10.1109/ICASSP.2008.4518795)
7. Rao P, Swarup P, Pasad A, Tulsiani H, Ghosh Das G: [Automatic assessment of reading with speech recognition technology](#). International Conference on Computers in Education. 2016. 1-3.

How to cite this article:

Dwivedi P, Jha A (April 23, 2026) System for Competency Assessment in Language Education (SCALE): A Computer-Mediated Assessment System for Measuring Proficiency in Indian Languages. Cureus J Comput Sci 3 : es44389-025-00055-x. DOI <https://doi.org/10.7759/s44389-025-00055-x>

System for Competency Assessment in Language Education (SCALE): A Computer-Mediated Assessment System for Measuring Proficiency in Indian Languages

8. Shermis MD, Burstein J, Higgins D, Zechner K: [Automated essay scoring: Writing assessment and instruction](#). International Encyclopedia of Education (Third Edition). 2010, 20-26. [10.1016/B978-0-08-044894-7.00233-5](#)
9. Wilson J, Roscoe R, Ahmed Y: [Automated formative writing assessment using a levels of language framework](#). Assessing Writing. 2017, 34:16-36. [10.1016/j.asw.2017.08.002](#)
10. Chalhoub-Deville M: [Language testing and technology: Past and future](#). Language Learning & Technology. 2001, 5:95-98. [10.64152/10125/25130](#)
11. Council for Cultural Co-operation, Education Committee, Modern Languages Division: [Common European Framework of Reference for Languages: Learning, Teaching, Assessment](#). Cambridge University Press, Cambridge, UK; 2001.
12. [ACTFL Professional Learning, Assessments](#). (2024). <https://www.actfl.org/news/updated-actfl-proficiency-guidelines-released>.
13. Pawlikowska-Smith G: [Canadian Language Benchmarks 2000: Theoretical Framework](#). Centre for Canadian Language Benchmarks, Ottawa, ON; 2002.
14. [Interagency Language Roundtable](#). Accessed: February 12, 2025: <https://www.govtilr.org/IRL%20History.htm>.
15. Dwivedi P: [Concept Based Continuum of Graded Syllabi \(CCGS\)](#). Dwivedi P (ed): NTS-India, CIIL, Mysore; 2016.
16. Stansfield CW: [Technology and Language Testing. A Collection of Papers from the Annual Colloquium on Language Testing Research \(7th, Princeton, New Jersey, April 6-9, 1985\)](#). Teachers of English to Speakers of Other Languages, Princeton, NJ; 1986.
17. Alderson JC: [Diagnosing Foreign Language Proficiency: The Interface Between Learning and Assessment](#). Continuum, London; 2005.
18. [Bulgarian Placement Test](#). Accessed: June 21, 2025: <https://www.17-minute-languages.com/en/learn-bulgarian/placement-test/>.
19. [Chinese Tests Service, HSK](#). Accessed: July 18, 2025: <https://www.chinesetest.cn/HSK>.
20. [German examinations A1-C2](#). Accessed: February 10, 2025: <https://www.goethe.de/en/spr/prf.html>.
21. [The Polish Language, Polish National Agency for Academic Exchange](#). Accessed: December 12, 2024: <https://nawa.gov.pl/en/the-polish-language/certification>.
22. [Universidad de Salamanca, Dele Exams, Deadlines And Exam Dates | Spanish Diplomas](#). Accessed: December 2, 2025: <https://www.dele.org/>.
23. Dwivedi P, Shankar RT, Meghana B, Sushaini H, Sudeep BR, Pooja MR: [Designing an adaptive question Bank and question paper generation management system](#). Advances in Artificial Intelligence and Data Engineering. AIDE 2019. Chiplunkar NN, Fukao T (ed): Springer, Singapore; 2021. 1133:965-973. [10.1007/978-981-15-3514-7_72](#)
24. Muhammaditya N, Hardjosoekarto S, Herwantoko O, Fany YG, Subangun MI: [Institutional divergence of digital item bank management in bureaucratic hybridization: An application of SSM based multi-method](#). Systemic Practice and Action Research. 2022, 35:527-553. [10.1007/s11213-021-09579-4](#)
25. Naik PG, Oza KS: [Version control system for auto generation of question papers](#). Smart Trends in Computing and Communications. SMART 2023. Senjyu T, So-In C, Joshi A (ed): Springer, Singapore; 2023. 645:371-384. [10.1007/978-981-99-0769-4_33](#)

How to cite this article:

Dwivedi P, Jha A (April 23, 2026) System for Competency Assessment in Language Education (SCALE): A Computer-Mediated Assessment System for Measuring Proficiency in Indian Languages. Cureus J Comput Sci 3 : es44389-025-00055-x. DOI <https://doi.org/10.7759/s44389-025-00055-x>