Innovations in English Proficiency Testing: A Focus on PTE Preparation Tools

[1] Akshara M S, [2]Dhaya Varghese ,[3]Roshini Kuriakose ,[4]Soorya Pradeep,[5]Krishnaveni V V
[1,2,3,4]Student,[5]Assistant Professor
[1,2,3,4,5]Department of Computer Science and Engineering
[1,2,3,4,5]College Of Engineering, Kidangoor, Kottayam, Kerala, India
[1]aksharams1111@gmail.com , [2]dhayasusan2003@gmail.com , [3]roshinik810@gmail.com ,
[4]pradeepsoorya73@gmail.com , [5]krishnaveni@ce-kgr.org

ABSTRACT: The Pearson Test of English (PTE) is a well-known test that checks how well non-native English speakers are at speaking, listening, writing and reading. Nowadays there are many digital resources, like as apps, to assist test-takers because the number of test takers is increasing on a daily basis. These apps are great because they offer specific practice, timed exercises, and feedback that feels like the actual test, making it easier for people to improve their skills step by step. But as online PTE preparation gets more popular, the future is moving towards advanced web platforms that are even more personal and interactive. These new platforms will use smart tech like advanced analytics and AI to give better feedback and create custom study plans that fit each user's needs. This means studying will be more effective and focused on what you really need to work on. The future of PTE preparation tools looks promising, with platforms that combine the ease of apps with the many features of web-based systems. These tools can be designed to meet the needs of all learners and help them succeed.

INDEX TERMS: Machine learning, Deep learning, Speech recognition

I.INTRODUCTION

In today's interconnected world, English has solidified its position as the global language. Its importance extends far beyond academic settings. Some key areas where English proficiency is vital including the areas of education, career opportunities, career advancement, communication and collaboration- Global Networking, business and trade, technology, innovation and scientific research.

The Pearson Test of English (PTE) Academic is a computer-based English language test widely recognized by universities and institutions globally. It assesses a test-taker's ability to communicate effectively in an academic setting. The PTE is designed to evaluate skills in speaking, listening, reading, and writing. One of the key advantages of the PTE is its computer-adaptive format, which ensures a fair and objective assessment of language abilities. This means that the difficulty level of the questions is adjusted based on the candidates performance, providing a more accurate evaluation.Another benefit of the PTE is the quick turnaround time for the results. Test-takers typically receive their scores within a few business days, allowing them to make timely decisions about their academic or professional plans. Its acceptance by a diverse range of institutions makes it a popular choice for individuals seeking to study or work abroad. To prepare for the PTE, a variety of resources and tools are available. These include practice tests, simulated exam conditions, and instant feedback to help test-takers identify their strengths and weaknesses. By utilizing these resources and practicing regularly, individuals can significantly improve their PTE scores and enhance their chances of achieving their academic and career goals. PTE Academic is a valuable tool for assessing English language proficiency. Its computer- adaptive format, quick results, and wide acceptance make it a reliable choice for people seeking to pursue higher education or career opportunities in English-speaking countries. By understanding the test format, practicing regularly, and utilizing available resources, test-takers can increase their confidence and achieve their desired PTE scores.

Emerging web platforms for PTE preparation integrate several advanced technologies to create an immersive and personalized learning experience. Artificial Intelligence (AI) plays a major role, enabling personalized feedback and adaptive learning, where exercises adjust based on user performance. Machine Learning (ML) algorithms further analyze user data, identifying patterns in strengths and weaknesses to create a customized study path. Natural Language Processing (NLP) is essential for automated scoring in the speaking and writing sections, allowing instant, detailed feedback on language accuracy, fluency, and coherence. Additionally, cloud computing supports these platforms, ensuring scalability, real-time access, and high-speed data processing, while data analytics provides insights into learner progress and potential areas of improvement. Furthermore, voice recognition technology offers advanced speaking practice by analyzing pronunciation and accent, helping candidates prepare for real-world speaking tasks.

II.LITERATURE SURVEY

For ease of study the literature is categorized based on the focus areas: A. Language Test Comparisons and Score Analysis B.Domain Relevance of PTE Academic for Australian Professional Bodies C.Test Scoring and Validity Studies D. Test Preparation and Second Language Acquisition E. Technology and Immersive Learning in Language Exams F. Technology-Assisted Language Learning and Educational Technology G. AI and Machine Learning in Education and Language Processing.

0.1 Language Test Comparisons and Score Analysis

A Comparative Study of IELTS and PTE Writing Scores [8]et al. examines differences in writing scores of both IELTS and PTE, in order to understand the factors that lead to variances in scoring. It investigates task design, rating criteria, automated scoring systems, and test administration as potential sources of disparity. This section of the survey focuses on capturing the experiences of the respondents regarding these factors and their perceived impact on test scores. For automated scoring versus human scoring, it explores how automated scoring in PTE assesses complex language skills, such as coherence and style, compared to human scoring in IELTS.

(a) Academic Task and Skills Relevance: The IELTS writing section focuses on tasks that reflect academic and professional writing scenarios, such as essay writing and data interpretation. These tasks assess critical thinking, argument structuring, and solution-oriented writing, which are essential for academic success and professional contexts. In contrast, the PTE tasks are designed to evaluate general linguistic ability through tasks like summarizing written text or essay writing, which are less specific to academic and professional requirements, potentially leading to a narrower assessment of skills.

(b) Proficiency Standards:IELTS employs human raters to assess writing, ensuring that the nuances of grammar, vocabulary, and argument development are carefully evaluated against high academic standards. The subjective nature of human scoring, however, may introduce variability between examiners, though rigorous training and standardization attempts to mitigate this. PTE's automated scoring system offers consistency and objectivity, but it can fail to capture elements like the test taker's creativity, tone, or persuasive writing style, which are critical for evaluating comprehensive proficiency.

(c)Authenticity of Task:IELTS writing tasks are designed to closely resemble real-world academic challenges, such as forming cohesive arguments and critically analyzing a topic. This authenticity prepares test takers for real academic tasks, making it highly relevant for those pursuing higher education or professional roles. On the other hand, PTE tasks prioritize linguistic accuracy and general writing structure, which, while useful, may lack the depth to fully replicate real-world academic or professional writing scenarios.

(d) Evaluation of Communication Skills:IELTS evaluates writing holistically, focusing on coherence, logical flow, vocabulary richness, and grammatical range to assess effective communication. This comprehensive approach captures the depth and quality of written communication but relies on human judgment, which can sometimes lead to inconsistencies. In contrast, PTE relies heavily on automated scoring, which objectively evaluates linguistic elements but may overlook subtleties like rhetorical strategies or nuanced argumentation, affecting the assessment of communication quality.

(e) Suggestions for Improved Standards:To improve IELTS, examiner training programs can focus on further standardizing scoring practices to reduce subjective variations and ensure fairness. PTE can enhance its evaluation process by incorporating a hybrid scoring model, combining automated scoring for objectivity and human assessment for nuanced evaluation. Both tests should provide detailed scoring criteria and task-specific feedback to test takers and educators to enhance transparency and preparation effectiveness.

Test takers	IELTS	PTE
Test taker 1	7.5	65
Test taker 2	7.0	73
Test taker 3	7.0	60
Test taker 4	6.5	75
Test taker 5	6.0	80
Test taker 6	5.5	81
Test taker 7	7.5	76
Test taker 8	6.0	64
Test taker 9	5.0	85
Test taker 10	5.5	77

Figure 1: Test taker's scores in IELTS and PTE [8]

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
ielts-score	.169	10	.200*	.930	10	.452
pte-score	.171	10	.200*	.940	10	.557

Figure 2: Tests of normality. [8]

	t	df	Sig. (2-tailed)	Mean difference	95% Confidence interval of the difference	
					Lower	Upper
Ielts-score	22.729	9	.000	6.3500	5.718	6.982
Pte-score	28.591	9	.000	73.600	67.78	79.42

Figure 3: T-test results. [8]

The figure 1 shows a noticeable diference in the scores of test takers between the IELTS and PTE exams. The Null hypothesis:""there is no statistically significant difference between the writing scores of PTE and IELTS, taken by the same individuals"[8]. As from figure 2 and figure 3, after doing t-test the p value was found to be less than 0.05 thus rejecting of null hypothesis. In conclusion there is a significant difference between PTE and IELTS scores when same individual is attempting. IELTS relies on human scoring, which assesses broader competencies like coherence, argumentation, and tone, whereas PTE's automated scoring system emphasizes measurable linguistic skills like grammar and vocabulary. This highlights the need to interpret the scores in the context of each test's unique scoring framework.

0.2 Domain Relevance of PTE Academic for Australian Professional Bodies

The research study reported by [6]et al. focuses on the relevants of PTE Academic Tasks to Professional Requirements such as:

(a) Alignment of PTE Academic tasks with skills relevant to Australian professional bodies. This section of the survey could gather perceptions from professionals on the applicability of specific PTE

tasks (e.g., summarizing, report writing) within their domains.

(b) Proficiency Standards and Performance Descriptors: The research examines proficiency levels defined by the Department of Home Affairs (DHA) and their relevance to professional compe- tencies. Survey questions in this section could explore the alignment between current PTE Academic standards and professional proficiency expectations. For example, the analysis could focus on evalu- ating the extent to which PTE Academic performance descriptors (e.g., Proficient, Competent) align with the language proficiency required in various fields and propose modifications to existing standards to better reflect the skills needed for professional practice.

(c) Authenticity of Tasks: The relevance of PTE Academic tasks, such as summarizing text and writing essays, to professional fields. Participants are invited to identify which PTE tasks closely mirror activities in their profession and to reflect on the importance of skills like summarizing, technical vocabulary, and data interpretation in their daily work.

(d) Evaluation of Communication Skills for Workplace Readiness: This study examines whether PTE tasks adequately measure the skills required for professional communication, especially for sectors such as healthcare, engineering, and accounting. The study explores whether PTE tasks ad- equately measure essential skills for professional communication, particularly in sectors like healthcare, engineering, and accounting. It investigates the effectiveness of PTE tasks in assessing readiness for professional communication and considers additional skills that could better reflect the demands of the language in the workplace.

(e) Skill-Specific Relevance; The panel noted that while many PTE skills were relevant, some tasks may require additional contextualization for profession-specific scenarios.While many PTE skills are recognized as relevant, some tasks may require further contextualization to align with profession-specific scenarios. This section delves into how PTE tasks could be adjusted to better meet professional communication needs, focusing on specific language skills, such as tone and technical accuracy, that may be underrepresented but are critical in professional settings.

(f) Suggestions for Improved Standards: To better support professional communication needs, the study seeks input on the refinement of the PTE performance descriptors. Focuses on proposed changes to make standards more relevant to specific fields and assesses the extent to which PTE proficiency levels enable new professionals to interact successfully in their industries.

0.3 Test Scoring and Validity Studies

The primary objective of the study is to develop more valid scoring criteria for integrated tasks in language assessment, specifically focusing on Reading-Writing (R-W) and Listening-Writing (L-W) summary tasks. The problem statement highlights the lack of scoring criteria that accurately capture the unique construct of integrated summary writing, particularly when automated scoring systems are used. This gap poses challenges in scoring validity, as current systems may overlook essential features of integrated writing, such as coherence, relevance of ideas, and source use. The study adopts a mixed-method approach, utilizing expert judgment, text analysis, and statistical analysis. It examines writing features that differentiate summary responses by 150 candidates at five proficiency levels on R-W and L-W tasks. Experts provided insights into salient features of summary writing across proficiency levels, and a detailed coding and statistical analysis were conducted to identify distinguishing features across these tasks. Major Findings and Contributions includes:

(a) Key Features for Proficiency Differentiation: There are several key features that vary across proficiency levels for R-W tasks, including relevance of ideas, paraphrasing skills, academic style, language control, coherence, and cohesion. However, the same features were less effective in differentiating proficiency for L-W tasks.

Organisation	Language proficiency requirement	PTE Academic score				
		Overall	Listening	Reading	Speaking	Writing
NMBA	Proficient	65	65	65	65	65
EA	Competent		50	50	50	50
AIMS	Proficient	65				
NAATI - Recognised Practising Interpreter and Certified Provisional Interpreter	Competent		51		59	
NAATI - Certified Interpreter and Certified Specialist Interpreter	Proficient		76	76	76	59
NAATI - Certified Conference Interpreter	Proficient		76	76	76	59
NAATI - Recognised Practising Translator	Competent			59		59
NAATI - Certified Translator	Proficient			76		76
NAATI - Certified Advanced Translator	Competent		59	76	59	76
CA ANZ	Proficient	65	65	65	65	65
CPA A (general skills assessment)	Proficient	65	65	65	65	65
CPA A (provisional skills assessment)	Competent	50	50	50	50	50
ASMIRT	Proficient	65	65	65	65	65
APC	Proficient	65	65	65	65	65

Figure 4: The language proficiency requirements of PTE Academic scores across language skills for the professional bodies. [6]

(b) Implications for Scoring Criteria: The research suggests expanding scoring categories to more accurately reflect integrated writing constructs, which could inform improvements in automated scoring systems. For L-W tasks, the study highlights the need for better descriptors that can effectively capture task-specific competencies.

(c) Potential for Automated Scoring: The findings support the development of machine-scored criteria that incorporate a broader set of features, moving beyond simple mechanical aspects to include nuanced factors of effective writing, such as source integration.

Limitations and Future Work Suggestions

Limitations: The study acknowledges that modality differences between R-W and L-W tasks limit the generalizability of findings across these formats. Additionally, the sample size, though stratified, may not capture the full diversity of candidate responses.

Future Work Suggestions: The authors recommend further research into scoring descriptors specifically for L-W tasks to address their unique demands. Expanding the sample size and including additional proficiency levels could also provide more comprehensive insights. [4] contributes to the body of knowledge by providing empirical evidence for feature-based scoring systems, potentially informing the development of more precise, fair, and reliable automated assessment tools in language testing.

The pie chart represents the study's main focus areas. Key Features for Proficiency Differentiation (R-W) accounts for 30 percent and highlights features like coherence, relevance, and language control that distinguish proficiency levels in Reading-Writing tasks. Implications for Scoring Criteria (L-W) also represent 30 percent, stressing the need for improved descriptors and scoring categories for Listening-Writing tasks. Potential for Automated Scoring makes up 20 percent, suggesting that machine-scored systems include nuanced factors like source integration. Lastly, Limitations and Future Work Suggestion constitute 20 percent, acknowledging challenges such as modality differences and recommending expanded research to enhance scoring systems.

0.4 Test Preparation and Second Language Acquisition

Language testing has become increasingly significant with globalization, which impacts education, career opportunities, and migration. High-stakes language tests like IELTS and TOEFL often require substantial preparation, which has led to a broad industry focused on test coaching and preparation strategies. However, the balance between effective language acquisition and test-oriented strategies remains contentious. This literature survey by [9] reviews recent studies on test preparation to an-

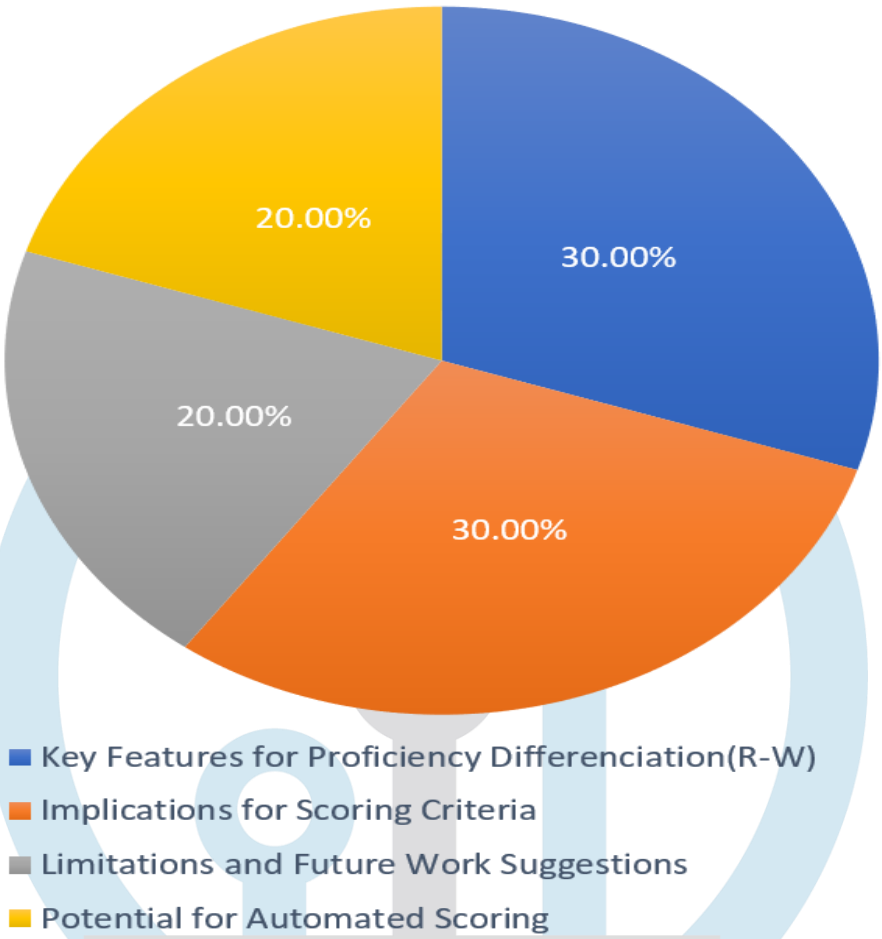

Figure 5: Summary of findings and contributions[9]

alyze methods, trends, gaps, and implications within this domain. It conducted a scoping review of 66 studies to examine the key characteristics of test preparation research, focusing on major themes, participant characteristics, and methodological patterns. The review emphasizes three main themes: perceptions, practices, and impacts. It was found that most studies were descriptive and often focused on learner perceptions and test preparation's impact on scores. Despite this, there is a notable lack of experimental studies and a scarcity of quantitative approaches.

The methodologies in test preparation research often favor surveys and thematic analysis, with descriptive designs predominating. [9] highlighted the need for more robust quantitative methods, as only a few studies employed statistical analyses such as factor analysis and structural equation modeling (SEM). The majority rely on thematic analysis, yet inconsistencies in reporting the analysis process were noted, suggesting a need for improved transparency.

There are three types of text preparation: construct-relevant skill development, familiarity with test formats, and strategies that might inflate scores without improving actual language proficiency. Studies frequently examine the impacts of preparation practices on test performance, but fewer assess the pedagogical implications or long-term language development outcomes. Comparative Analysis and Discussion While the scope of existing research is broad, the review reveals a heavy concentration on English language tests, primarily IELTS and TOEFL, with significant gaps in research on other languages and contexts. Additionally, the prevalence of descriptive over experimental research limits insights into causality in test preparation practices. Methodologically, the field may benefit from incorporating more diverse, experimental designs and including participants beyond learners and teachers, such as administrators and policymakers, to broaden understanding of test preparation's washback effects on teaching and policy. This literature survey highlights that while test preparation research has evolved, there remain considerable gaps, particularly in experimental approaches and broader language coverage. Future studies should expand the focus to include diverse language tests and explore advanced statistical techniques to enhance understanding of preparation's effects on both test performance and language acquisition.

0.5 Technology and Immersive Learning in Language Exams

As advancement in technology, how we learn language has changed dramatically, especially with the advent of mobile-assisted learning platforms and immersive learning technologies. The purpose of this literature review is to investigate current developments in the use of mobile-assisted pronouncing aids, augmented reality, and virtual reality in English language teaching (ELT). Teachers and researchers can design more successful language learning experiences by having a better understanding of the effects, constraints, and possibilities of these technologies. Technologies for Immersion Learning in ELT The relevance of immersive technologies in ELT is examined in systematic review , which highlights the differences between VR and AR. AR's visual engagement aids in vocabulary development and VR's immersive environment helps in language-speaking practice. According to this survey [1], 2017 marked a shift in interest in these technologies, which is associated with mobile accessibility. However, there was also documentation of problems including high cost and technical restrictions.

The review of mobile-assisted language learning [10] focuses on pronunciation using mobile-assisted language learning (MALL). Pronunciation apps have gained popularity among EFL learners due to the increase in mobile usage, particularly on smartphones. [10] emphasizes how user engagement and pronunciation have increased due to resources such as English Pronunciation Tutor. The study also highlights the need for more educational assistance to ensure sustained learning outcomes even though MALL offers an appealing approach.

0.6 Technology-Assisted Language Learning and Educational Technology

Artificial intelligence (AI) and technology integration in the classroom are now essential for improving programming and language learning. AI based tutoring programs and technology-assisted language learning (TALL) offer personalized learning experiences that promote interest and skill improvement. To better understand the trends, approaches, and gaps in research on TALL systems and AI tutoring, this overview review recent studies. This survey determines the main contributions, constraints and possible future directions in these fields by examining recent reviews systematically and case studies.

(a) Technology-Assisted Language Learning (TALL) A comprehensive study of TALL was carried out by [3], with an emphasis on the benefits, language proficiency, and forms of technology. They found that TALL systems, which mostly target English language learners, frequently use intelligent systems, mobile learning, and game-based feedback mechanisms. Although there were some drawbacks, like learning anxiety and difficulties with technology literacy, the main advantages were increased motivation, re- tention, and collaborative learning Technology-Assisted-Language learning. Although highlighting a deficit in the analysis of long-term language proficiency progress, their work adds to our understanding of the influence of TALL.

(b) AI-Tutoring in the Teaching of Programming The usage of GPT-3.5-Turbo as an AI-Tutor within an Automated Programming Assessment System (APAS) was the aim of [7] investigation into AI-based tutoring in programming. Their study categorized candidates based on interaction patterns, showing suitable results in prompt feedback and user involvement. students' worries about an increased dependence on AI were among the limits, though and they suggested enhancements to interactive design and personalized feedback. While this study contribute to AI in programming education, further research is needed to fully understand how AI affects programming proficiency.

(c) Discussion and Comparative Analysis While AI-Tutoring focuses on programming and TALL research on language learning, both studies highlight the significance of technology in personalized learning. The limited use of sophisticated experimental techniques, which could offer more deep understanding into the effectiveness of these systems, is a significant gap in both fields. A variety of devices are useful for TALL systems but more complex interaction models are needed for AI tutoring to reduce dependence and improve learning retention. Future studies should examine a range of learner demographics, going beyond the usual.

(d) Conclusion and future directions: This review underline how well technology may improve learning outcomes in language and program learning. Although current research showcases the potential of these systems, more research is needed to refine AI feedback mechanisms, broaden participant di-

versity and examine long-term impacts. Filling these gaps could reveal insights for the development of context-sensitive adaptive educational tools that bridge the gap between technology and effective learning experiences. However, limitations included generic responses and students' concerns about over-reliance on AI, suggesting improvements in personalized feedback and interactive design.

0.7 AI and Machine Learning in Education and Language Processing

(a) Multimodal Dialogue Situations in Educational Environments Goal and Approach: In order to identify discussion situations,[5] presented a revolutionary multimodal dialogue recognition framework that integrates linguistic, visual, and auditory data. This method classifies context cues, such as speech styles, locations, activities, and participant relationships, by utilizing Gated Recurrent Units (GRU) and Bidirectional Encoder Representations from Transformers (BERT).

Results and Inputs: The results of the study show that integrating several modalities greatly enhances context recognition in instructional conversations, leading to high classification accuracy for dialogue scenarios. This multimodal method is essential in educational contexts where real-time lesson adaptation depends on an awareness of conversational context, student involvement levels, and interac- tion quality. The study highlights the necessity of dialogue systems that are flexible enough to decipher situational clues to provide tailored feedback, improve understanding, and react appropriately in light of the circumstances.

Restrictions and Prospects for the Future: Uncertainty in categorizing overlapping circumstances presented difficulties for the study, which is a drawback in educational applications where context and degrees of engagement can vary. To improve model adaptation in educational dialogue systems, future research can focus on improving situation labels and adding more contextual information, including participant age and prior learning experiences.

(b) Machine Learning Models for Adaptive Performance Prediction Goal and Approach: Uses "Learning Coefficients" obtained from trajectory-based adaptive assessments of [2] to meet the requirement for real-time, adaptive performance prediction in educational contexts. These coefficients improve the responsiveness of conventional machine learning by acting as dynamic predictors of student achievement through ongoing evaluation of student answers.models in learning environments. In order to analyze prediction accuracy using learning coefficients, this study used a variety of regression-based machine learning models, such as support vector regression, decision trees, random forests, and linear regression.

Results and Inputs: According to the results, the learning coefficients helped the prediction models reach high precision, allowing teachers to continuously assess the comprehension of their students. By providing a framework for real-time feedback, the study enables students to spot and fill in knowledge gaps early on in the learning process. This work demonstrates the potential of dynamic evaluation tools in the processing of educational language by offering personalized recommendations for improvement and using adaptive feedback to reinforce comprehension in real time.

Restrictions and Prospects for the Future: The study's conclusions were not as broadly applicable as they could have been due to its small dataset. The use of learning coefficients could be expanded and validated by integrating more varied student profiles and across different educational levels. In order to guarantee that these predictive tools are generally relevant, future studies may also examine how well adaptive models function in a variety of subjects and learning contexts.

The above table represents a comparison of various technologies or methodologies across ten papers. A checkmark indicates the presence or relevance of a specific technology in each paper.

III.CONCLUSION

In summary, the Pearson Test of English (PTE) has become a vital tool for assessing English proficiency among non-native speakers. The paper explores the strengths of PTE preparation tools, particularly mobile applications that provide accessible, flexible, and targeted practice. Although these tools support candidates in improving their performance, there is potential for future platforms to incorporate advanced technologies such as artificial intelligence and adaptive learning. These technologies can im-

	Machine Learning	Speech Recognition	Data Processing Tools
[1]	✓		
[2]	✓		✓
[3]	✓	✓	
[4]		✓	
[5]	✓	✓	✓
[6]			✓
[7]	✓	✓	✓
[8]	✓		✓
[9]			✓
[10]	✓	✓	

Figure 6: Technologies used

prove personalization,enabling a tailored learning path that meets individual needs and challenges. As demand for PTE and other English proficiency tests grows, the evolution of these digital resources promises an integrated ecosystem that combines convenience with robust, real-time support for test-takers. This convergence of technology and education could set new standards for language learning, helping more candidates reach their desired proficiency levels effectively and confidently.

References

[1] H. K. Altun and J. Lee. Immersive learning technologies in english language teaching: A systematic review. *Educational Technology International*, 21(2):155–191, 2020.

[2] P. Asthana, S. Mishra, N. Gupta, M. Derawi, and A. Kumar. Prediction of student's performance with learning coefficients using regression based machine learning models. *IEEE Access*, 2023.

[3] H. Buddha, L. Shuib, N. Idris, and C. I. Eke. Technology-assisted language learning systems: A systematic literature review. *IEEE Access*, 2024.

[4] S. Chan and L. May. Towards more valid scoring criteria for integrated reading-writing and listening-writing summary tasks. *Language Testing*, 40(2):410–439, 2023.

[5] Y. Chiba and R. Higashinaka. Dialogue situation recognition in everyday conversation from audio, visual, and linguistic information. *IEEE Access*, 2023.

[6] R. Clesham, S. Hughes, and L. Miller. Examining the domain relevance of pte academic for australian professional bodies. 2023.

[7] E. Frankford, C. Sauerwein, P. Bassner, S. Krusche, and R. Breu. Ai-tutoring in software engineering education. In *Proceedings of the 46th International Conference on Software Engineering: Software Engineering Education and Training*, pages 309–319, 2024.

[8] H. Ghaemi and Z. Khorsand. A comparative study of writing scores in ielts and pte: An investigation into the factors leading to disparity in scores. *African Educational Research Journal*, 12(1):6–20, 2024.

[9] S. He, A.-M. Sénécal, L. Stansfield, and R. Suvorov. A scoping review of research on second language test preparation. *Language Testing*, page 02655322241249754, 2024.

[10] R. Metruk. Mobile-assisted language learning and pronunciation instruction: A systematic literature review. *Education and Information Technologies*, pages 1–28, 2024.