Beyond Paper and Pencil: Evaluating the Effectiveness and User Perception of a Digital Reading Proficiency Assessment Platform

Ivy M. Tarun
Isabela State University, Cabagan, Philippines

***Abstract*—This study investigated the effectiveness of a digital reading proficiency assessment platform compared with traditional paper-and-pencil methods and examined users' perceptions of its performance. It employed a quasi-experimental pretest–posttest control group design involving two comparable groups of secondary school students: a control group assessed using conventional reading assessment procedures and an experimental group assessed using the digital platform for reading proficiency assessment. Pre-test and post-test scores in linear and non-linear reading tasks were analyzed using paired and independent sample t-tests to determine differences in reading proficiency gains between the two groups. To evaluate user perception, survey questionnaires grounded in the ISO/IEC 25010 software quality standards were administered to students and teachers, focusing on functional suitability, usability, security, and maintainability. Results revealed that the experimental group achieved significantly higher reading proficiency gains than the control group in both assessment types, with statistical analyses confirming that the improvements were significant. In contrast, the control group showed minimal or no meaningful improvement. User perception findings indicated a high level of satisfaction across all evaluated quality dimensions, suggesting that the digital assessment platform was perceived as reliable, user-friendly, secure, and adaptable to instructional needs. The results provide empirical evidence that digital reading proficiency assessment is both more effective and more positively perceived than traditional methods. The study concludes that integrating digital assessment platforms in reading evaluation can enhance assessment accuracy, efficiency, and user acceptance, offering a viable and evidence-based alternative for improving reading assessment practices in secondary education.**

Keywords*—Digital reading assessment; reading proficiency; quasi-experimental design; user perception; software quality evaluation

I. Introduction

In today's rapidly evolving education system, technology has become an indispensable tool for improving and streamlining the delivery of educational services. It continues to make teaching and learning more efficient, accessible, and responsive to the needs of students. Beyond convenience, it offers innovative ways to assess and enhance students' academic skills, particularly reading comprehension, which the authors in [1] identify as a foundation for future learning. Despite its critical role, reading proficiency among Filipino students remains a pressing concern. Results from the Program for International Student Assessment reveal that only 24% of Filipino students achieved at least Level 2 proficiency, far below the Organization for Economic Cooperation and Development (OECD) average of 74% [2]. This alarming gap underscores the urgent need for effective and sustainable literacy interventions.

In response, the Department of Education (DepEd) launched the Catch-Up Fridays initiative in 2024, integrating the Drop Everything And Read (DEAR) program to strengthen students' reading skills through structured and regular activities. The implementation of this initiative in the Philippines relies heavily on manual processes such as downloading and sorting reading materials, conducting fluency checks, administering quizzes, and encoding results. While these practices support literacy development, they are time-consuming and susceptible to delays, errors, and inconsistencies, which can limit their overall impact.

Recognizing these challenges, a web and mobile-based application called "Platform for Efficient Evaluation of Reading Proficiency", or PEER, was built by a team of developers to support more efficient and consistent reading proficiency assessment. With its web platform, it features role-based access control, reading material management, quiz comprehension management, formula-based reading proficiency assessment, and weekly performance comparison. On the other hand, its mobile app provides a student portal, reading fluency and quiz comprehension and a historical record of reading proficiency levels. By automating key assessment tasks and enabling weekly progress tracking, the platform provides educators with timely and actionable insights into students' reading performance. In doing so, PEER empowers teachers to deliver targeted interventions, strengthen instructional practices, and more effectively address learning gaps.

To support evidence-based decision-making for the potential school-wide implementation of PEER, this study was undertaken to provide a more comprehensive analysis of its effectiveness in comparison with traditional paper-and-pencil assessment methods as well as users' perceptions of its performance. This study promotes inclusive and equitable educational practices, which ultimately highlight how thoughtfully integrated technology can enhance assessment processes, support educators, and contribute to improved literacy outcomes in secondary education.

II. Related Work

Digital technologies have increasingly been used to support reading instruction and assessment, particularly because they

can automate scoring, standardize procedures, and generate timely progress data that teachers can actually use. In large-scale research, digital reading performance has been linked not only to students' skills but also to classroom conditions and how instruction is experienced by learners. Using PISA data across multiple OECD countries [3] reveals that students' perceptions of instruction quality were significantly associated with digital reading performance, suggesting that technology-based reading outcomes are shaped by both learner factors and instructional context.

At the classroom level, empirical studies show that well-designed web-based reading systems can improve comprehension outcomes. The authors in [4] developed a web-based collaborative reading annotation system with gamification and reported gains in learners' reading comprehension compared with more conventional approaches, highlighting how structured digital interaction (e.g., annotation, feedback, and guided engagement) can support deeper processing of texts. Complementing this, [5] emphasized the "high potential" of computer-based reading assessment, noting that computerized tools can assess multiple components of reading and provide more fine-grained information than single-shot paper tests, including for secondary learners. In the study by [6], it was noted that digital tools can effectively boost adolescents' interest in reading, while also helping improve their reading skills and test performance.

However, research also cautions that "digital" is not automatically better unless the assessment experience is carefully designed. The meta-analysis in [7], comparing paper versus digital media, reports that comprehension outcomes can differ by medium and conditions, reinforcing the idea that digital assessments should be designed to support focused reading and comprehension, rather than simply delivering texts on screens. Related evidence from digital-text interventions suggests that embedded questions and feedback can improve comprehension and self-regulation while reading digitally, especially when learners are supported in navigating and integrating information.

Another consistent theme in the literature is the value of progress-monitoring and data-driven instruction. In a review of curriculum-based measurement (CBM), [8] concluded that frequent, systematic measurement can improve student achievement when teachers use the data to adjust instruction. Similarly, [9] found that learning progress assessment combined with goal setting can positively affect reading achievement and motivational outcomes, strengthening the case for platforms that support continuous monitoring rather than episodic testing.

Finally, adoption and sustained use depend heavily on user acceptance and perceived quality. The Technology Acceptance Model (TAM) established perceived usefulness and perceived ease of use as central drivers of user acceptance, an influential lens for understanding teacher and student uptake of educational platforms [10]. For system-level evaluation, many education technology studies map user feedback to recognized software quality frameworks; ISO/IEC 25010 provides widely used quality characteristics (e.g., functional suitability, usability, security, maintainability) for specifying and evaluating software systems [11]. Together, these works support the need to test digital reading assessment platforms not only for effectiveness (learning gains) but also for user perception and quality, directly aligning with studies that compare outcomes against traditional methods and assess perceived performance using quality standards.

Uncertainties remain about the effects of using digital tools in reading beyond paper and pen. In [12], the authors emphasize the importance of examining how the use of digital devices can affect cognitive processes, as well as how different digital features may either support or hinder reading comprehension.

III. Methodology

A. Procedures

This study employed a quasi-experimental pretest–posttest control group design to examine the effectiveness of the PEER-based assessment platform in one of the secondary schools in the northern part of the Philippines, which is the Delfin Albano High School (DAHS). The system evaluated in this study was developed by a separate development team and served as the intervention platform.

Two comparable groups of secondary school students participated in the study: one group utilized the PEER platform using their own mobile devices, while the other group continued with the school's existing paper-and-pencil reading assessment practices. Intact classes were assigned to either the control or experimental group, as random assignment of individual students was not feasible in the school setting. Teachers in the control group conducted reading assessments using their usual methods, whereas teachers in the experimental group implemented the PEER platform for reading proficiency assessment after their training with the developers.

A pretest was administered to both groups to establish baseline reading proficiency levels. This included linear tasks where students read text-based materials and then answered related comprehension questions or quizzes. It also included non-linear tasks where students worked with visuals like graphs and charts and answered questions that required them to interpret and analyze these materials. After a one-month implementation period, a posttest was conducted using the same types of activities to determine changes in reading performance.

To complement the quantitative findings, a survey was conducted among students, class advisers, and teachers who utilized the PEER platform to evaluate its perceived software quality. Respondents for the software evaluation were selected through convenience sampling. The assessment of user perception was guided by the ISO/IEC 25010 software quality model, focusing on functional suitability, usability, security, and maintainability.

Prior to data collection, respondents were informed of the purpose of the study and assured that their participation was voluntary. They were also informed of their right to withdraw at any stage of the study and that all responses would be treated with confidentiality and used solely for research purposes.

For data analysis, paired and independent samples t-tests were used to compare pretest and posttest scores within and between the control and experimental groups to determine whether observed differences in reading proficiency gains were

statistically significant, while Cohen's *d* was also used to determine the effect size of Reading Proficiency Gains. Survey data were analyzed using the weighted mean to describe perceived software quality. A two-factor ANOVA without replication was conducted to examine further the differences across respondents' ratings and software quality criteria.

B. Platform Under Study

The digital reading proficiency assessment platform called PEER served as the intervention platform that was built by a team of developers following the Agile methodology. Its development unfolded as a deliberate and iterative process focused on improving how reading proficiency is assessed in the locale of the study. Their work began by clarifying the goals of the system and understanding how different users, such as teachers, administrators, and students, would engage with both the web and mobile applications. Through consultations and use case analysis, essential features were identified and prioritized to support efficient assessment, progress monitoring, and meaningful feedback.

With these goals in place, attention shifted to designing interfaces that were simple, intuitive, and responsive to users' needs. Early layouts and interaction flows were refined to ensure that the web and mobile applications of the platform offered a smooth and accessible experience. System development then brought these designs to life by integrating assessment logic, data storage, and synchronized access between the web and mobile applications. Continuous testing was carried out throughout this phase to ensure stability, accuracy, and ease of use. Fig. 1 shows the Admin Dashboard of the PEER web application [13]. This serves as the main control center where administrators manage key academic data such as sections, faculty profiles, student records, enrollment, and reading materials. It also provides an overview of system statistics, including the total number of enrolled students and faculty members.

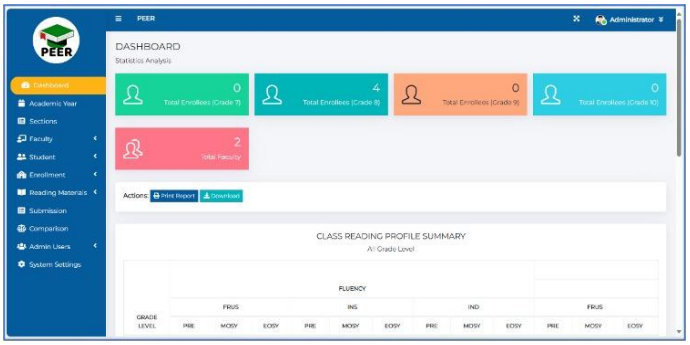

Fig. 1. Admin Dashboard of the PEER web App.

In addition, the dashboard features a class reading profile summary that allows users to quickly view and monitor student performance at a glance. Fig. 2 shows the Login Page of the PEER mobile application [13], which serves as the entry point for students before accessing the system's features. Through this interface, users are required to authenticate their accounts.

Fig. 3 shows the Reading Material Management, where teachers can upload and organize reading materials for student access, and Fig. 4 shows the Quiz Comprehension Management, which allows teachers to create and manage quizzes based on the selected reading material [13].

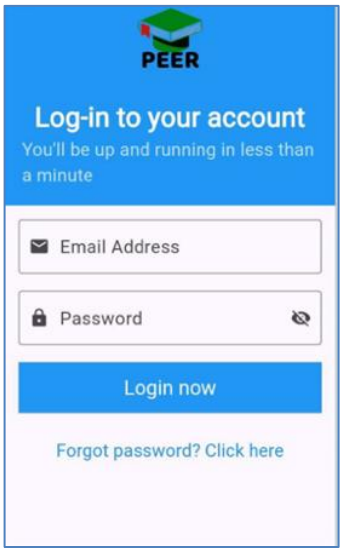

Fig. 2. Login Page of the PEER web app.

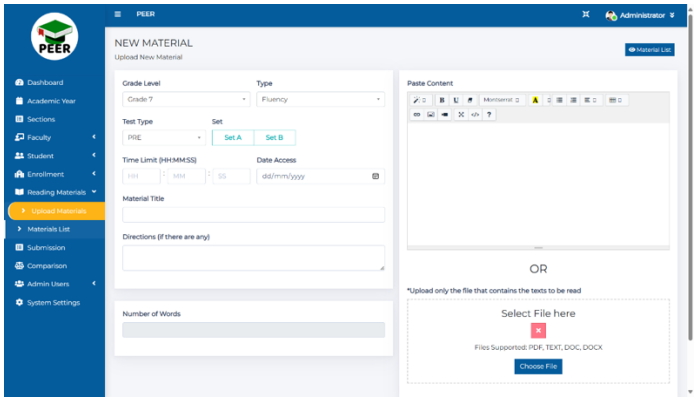

Fig. 3. Reading Material Management.

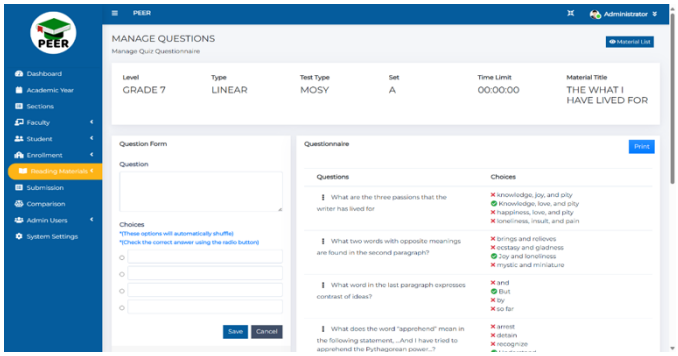

Fig. 4. Quiz Comprehension Management.

Fig. 5 shows the home page of the mobile application [13], including its simple and straightforward navigation menu that allows students to easily access the system's main features. Fig. 6, on the other hand, presents a sample recorded reading activity of a student, along with the quiz interface [13] where learners answer comprehension questions based on the assigned reading material.

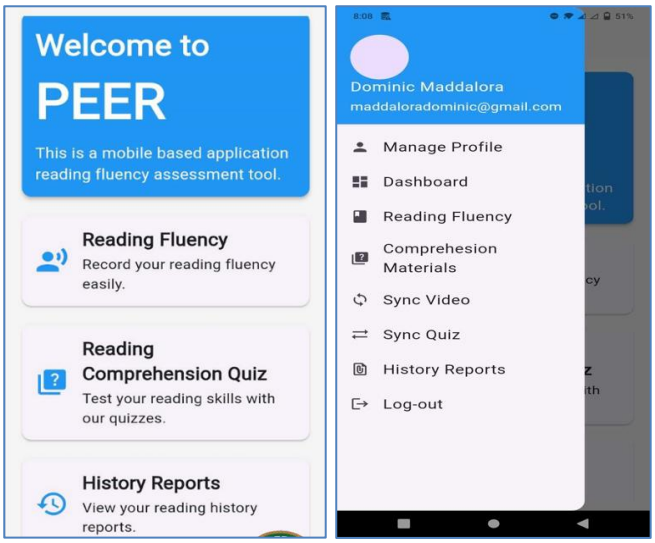

Fig. 5. Student portal.

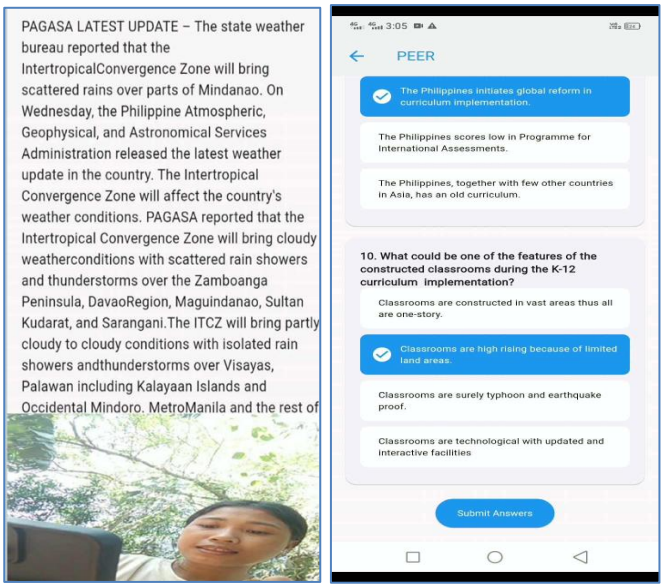

Fig. 6. Recorded student reading activity and quiz.

Following successful testing, the PEER platform was implemented by the developers in a public secondary school in Isabela, Philippines, and made available to first- and second-year high school students.

IV. Results

This section presents and discusses the findings of the platform's evaluation, focusing on a more constructive analysis of its effectiveness and user perception. The results are intended to provide empirical evidence to inform data-driven decisions regarding the potential adoption and wider implementation of PEER at the school level.

To examine the effectiveness of the PEER platform in comparison with traditional assessment methods, a t-test was employed as the primary statistical tool. Prior to conducting the t-test, the distribution of the data was assessed to ensure that the assumptions for parametric analysis were met. This assessment involved examining the skewness and kurtosis values of the scores.

As shown in Table I and Table II, the skewness values are close to zero, indicating that the score distributions are generally symmetric and did not exhibit substantial skewness. This suggests that the data were reasonably balanced around the mean. In addition, the kurtosis values are within acceptable ranges, further supporting the assumption that the data approximated a normal distribution. Taken together, these results indicate that the scores satisfied the normality assumption required for the application of t-tests.

TABLE I. Pre-Test and Post-Test Scores within Groups Across Linear and Non-Linear for the Control Group.

Measure of Data Distribution	Control Group			
	Pre-test Score		Post-test Score	
	Linear	Non-Linear	Linear	Non-Linear
Skewness	-0.041	-0.081	0.004	-0.022
Kurtosis	-0.174	-0.628	-0.899	-0.708

TABLE II. Pre-Test and Post-Test Scores within Groups Across Linear and Non-Linear for the Experimental Group.

Measure of Data Distribution	Experimental Group			
	Pre-test Score		Post-test Score	
	Linear	Non-Linear	Linear	Non-Linear
Skewness	-0.492	2.370	-0.674	-1.537
Kurtosis	-0.966	6.253	-0.447	0.961

A. Test of Difference Between the Pre- and Post-Test Scores Within Groups

After establishing that the data met the assumption of normality, a paired t-test comparing pre-test and post-test scores for both linear and non-linear reading assessments was computed.

For the control group, results presented in Table III indicate little to no improvement in reading performance. Mean scores in linear reading slightly declined from 4.067 in the pre-test to 3.767 in the post-test, with post-test scores showing slightly less variability (variance of 1.22 compared to 1.86 in the pre-test). The weak positive correlation between pre-test and post-test scores ($r = 0.217$), together with a non-significant p-value ($p = 0.300$), suggests that the observed changes were not statistically meaningful. These findings indicate that the traditional assessment approach did not lead to measurable improvements in students' reading performance. On the other hand, non-linear reading scores in the control group increased modestly from a mean of 4.400 to 4.933. Although the change reached statistical significance ($t = –2.570$, $p = 0.016$), the gain was small, indicating limited improvement using the traditional approach.

In contrast, the experimental group, as shown in Table IV, demonstrated substantial and statistically significant gains after using the PEER platform. Mean linear reading scores increased markedly from 4.567 to 8.033, while non-linear reading scores rose from 5.167 to 9.033. In addition to higher post-test means, the reduced variance in post-test scores reflects more consistent

performance among students. Although the correlations between pre-test and post-test scores were weak (linear: r = 0.085; non-linear: r = 0.181), the extremely small p-values ($p = 2.099 \times 10^{-9}$ for linear and $p = 9.004 \times 10^{-12}$ for non-linear) clearly indicate that these improvements were statistically significant.

TABLE III. RESULTS OF THE PAIRED T-TEST FOR PRE-TEST AND POST-TEST SCORES WITHIN CONTROL GROUPS.

	CONTROL GROUP			
	Linear		Non-Linear	
	Pre	Post	Pre	Post
Mean	4.067	3.767	4.400	4.933
Variance	1.857	1.220	1.628	1.237
Pearson Correlation	0.217		0.554	
t-stat	1.055		-2.570	
P(T<=t) one-tail	0.150		0.008	
t Critical one-tail	1.699		1.699	
P(T<=t) two-tail	0.300		0.016	
t Critical two-tail	2.045		2.045	

TABLE IV. RESULTS OF THE PAIRED T-TEST FOR PRE-TEST AND POST-TEST SCORES WITHIN EXPERIMENTAL GROUPS.

	EXPERIMENTAL GROUP			
	Linear		Non-Linear	
	Pre	Post	Pre	Post
Mean	4.567	8.033	5.167	9.033
Variance	3.357	2.033	1.868	2.723
Pearson Correlation	0.085		0.181	
t-stat	-8.537		-10.901	
P(T<=t) one-tail	1.049E-09		4.502E-12	
t Critical one-tail	1.699		1.699	
P(T<=t) two-tail	2.099E-09		9.004E-12	
t Critical two-tail	2.045		2.045	

Taken together, the results indicate that students assessed using the PEER platform achieved significantly greater gains in both linear and non-linear reading proficiency than those assessed using traditional methods, highlighting the platform's effectiveness as a digital assessment tool.

B. Test of Difference on the Reading Proficiency Gains Between the Control and Experimental Groups

Table V presents the mean gain scores in linear and non-linear reading assessments for both the control and experimental groups. The results show a clear contrast between the two groups in terms of reading proficiency gains. For the control group, the mean gain score in linear reading was -0.30, indicating a slight decline in performance from pre-test to post-test. In non-linear reading, the control group exhibited only a modest improvement, with a mean gain score of 0.53. These findings suggest that the traditional paper-and-pencil assessment approach resulted in little to no meaningful improvement in students' reading proficiency over the duration of the study.

In contrast, the experimental group demonstrated substantial gains in both reading assessment types. The mean gain score for linear reading was 3.47, while non-linear reading showed an even higher gain of 3.87. Independent samples t-test results revealed that the differences in gain scores between the control and experimental groups were statistically significant for both linear ($p = 2.88737 \times 10^{-10}$) and non-linear reading ($p = 3.99054 \times 10^{-11}$). These extremely small p-values provide strong evidence that the observed improvements in the experimental group were not due to chance but were associated with the use of the PEER platform.

TABLE V. MEAN GAIN SCORES FOR CONTROL AND EXPERIMENTAL GROUP.

	CONTROL GROUP		EXPERIMENTAL GROUP	
	Linear	Non-Linear	Linear	Non-Linear
MEAN GAIN SCORE	-0.30	0.53	3.47	3.87
p-value for Linear	2.88737E-10			
p-value for Non-Linear	3.99054E-11			

n=30

To determine the magnitude of the differences between groups, Cohen's *d* was computed using gain scores. Results revealed a very large effect size for both linear reading (d = 1.95) and non-linear reading (d = 2.02). These findings indicate that the use of the PEER platform produced a substantial and practically significant improvement in reading proficiency compared to traditional assessment methods (see Table VI).

TABLE VI. EFFECT SIZE (COHEN'S D) OF READING PROFICIENCY GAINS BETWEEN CONTROL AND EXPERIMENTAL GROUPS

Reading Assessment Type	Mean Gain		Cohen's d	Interpretation
	Control	Experimental		
Linear	-0.30	3.47	1.95	Very Large Effect
Non-Linear	0.53	3.87	2.02	Very Large Effect

The findings imply that the PEER-based assessment approach was significantly more effective than traditional methods in improving students' reading proficiency. The larger gains observed in the experimental group may be attributed to the platform's structured, consistent, and technology-supported assessment process, which allowed for timely monitoring of student performance and more focused engagement with reading tasks. Moreover, the greater gains in non-linear reading suggest that digital assessment environments may be particularly effective in supporting skills related to interpreting and integrating information from non-sequential texts. These results reinforce the value of integrating digital assessment platforms as part of reading proficiency programs to achieve stronger and more consistent learning outcomes.

C. Evaluation on the Performance of PEER

An evaluation was conducted using a survey questionnaire among the respondents to determine their perception with

respect to the functional suitability, usability, security and maintainability of the developed platform.

For the web application, all quality criteria received mean ratings above 4.50, corresponding to a descriptive rating of *Strongly Agree*. Functional suitability obtained a mean score of 4.72, suggesting that users perceived the system's features as well-aligned with their assessment needs. Usability also received a high mean of 4.69, indicating that users found the interface intuitive and easy to navigate. Similarly, security and maintainability were rated highly, with mean scores of 4.73 and 4.67 respectively, reflecting user confidence in data protection measures and the system's ability to support updates and modifications without difficulty.

Comparable results were observed for the mobile application, which also achieved *Strongly Agree* ratings across all criteria. Functional suitability (M = 4.57) and usability (M = 4.56) were particularly well rated, highlighting the effectiveness of the mobile app in supporting reading assessment tasks and providing a user-friendly experience. Security and maintainability each recorded mean scores of 4.52 and 4.55, respectively, indicating that users felt their data were secure and that the application was reliable and adaptable to future needs.

The overall mean scores of 4.70 and 4.55 for web and mobile platforms, respectively, underscore the consistency of user satisfaction across deployment environments. These findings suggest that the PEER platform meets high software quality standards and is perceived as reliable, secure, and easy to use by its users. The consistently strong ratings across all dimensions support the platform's suitability for broader implementation and reinforce its potential as an effective digital tool for reading proficiency assessment (see Table VII).

TABLE VII. PERCEIVED PERFORMANCE EVALUATION OF PEER

SW Quality Criteria	Web App		Mobile App	
	Mean	SD	Mean	SD
Functional Suitability	4.72	0.30	4.57	0.45
Usability	4.69	0.30	4.56	0.50
Security	4.73	0.24	4.52	0.47
Maintainability	4.67	0.31	4.55	0.48
Overall Mean	**4.70**		**4.55**	**0.49**

n=50

A two-factor ANOVA without replication was conducted to examine differences across respondents' ratings and software quality criteria. The computed values on Table VIII and Table IX revealed a statistically significant effect of respondents (for web: $F = 8.92, p < 0.001$ and for mobile: $F = 28.50, p < 0.001$), indicating that there were significant differences in the ratings provided by respondents. This suggests variability in how individuals evaluated the system. In contrast, the effect of software quality criteria (Functional Suitability, Usability, Security, and Maintainability) was not statistically significant, (for web: $F = 1.63, p = 0.202$ and for mobile: $F = 0.80, p = 0.45$). The findings indicate that while respondent evaluations on the web and mobile applications significantly varied, there were no statistically significant differences among the assessed software quality criteria. This suggests that the system demonstrates a consistent level of quality across all evaluated dimensions.

TABLE VIII. TWO-FACTOR ANOVA (WEB APP)

Source of Variation	SS	df	MS	F	P-value	F crit
Respondents	9.52	49	0.19	8.92	5.02E-20	1.48
SW Quality	0.07	2	0.04	1.63	0.201699	3.09
Error	2.14	98	0.02			
Total	11.73	149				

TABLE IX. TWO-FACTOR ANOVA (MOBILE APP)

Source of Variation	SS	df	MS	F	P-value	F crit
Respondents	32.25	49	0.656	28.50	1.49E-40	1.48
SW Quality	0.037	2	0.02	0.80	0.452383	3.09
Error	2.26	98	0.02			
Total	34.55	149				

V. DISCUSSION

This study sought to examine whether a digital reading assessment platform could improve students' reading proficiency compared with conventional paper-based assessment practices. The results provide encouraging evidence that the PEER platform contributed to measurable improvements in students' reading performance while also receiving strong acceptance from users. When viewed together, the statistical findings and user evaluations suggest that technology-supported assessment systems may play a meaningful role in strengthening reading evaluation practices in schools. However, it is worth noting that the study only involved a relatively small sample size, which may limit how far the findings can be applied to a wider population. Because of this, future studies are encouraged to test the platform with a larger and more diverse group of participants to further strengthen and confirm its effectiveness.

The results of the pre-test and post-test comparison revealed a clear contrast between the control and experimental groups. Students who participated in the traditional paper-and-pencil assessment showed little change in their performance across the two testing periods. In linear reading, the slight decrease in the mean score suggests that repeated exposure to similar reading tasks alone may not necessarily lead to improvement. Although the control group demonstrated a modest increase in non-linear reading scores, the magnitude of the improvement remained relatively small. This outcome reflects a limitation often discussed in assessment literature, where conventional testing approaches tend to emphasize measurement rather than actively supporting the learning process [14].

A different pattern emerged among students who used the PEER platform. Their post-test scores increased substantially in both linear and non-linear reading assessments, indicating that the digital platform may have created a more supportive environment for completing reading tasks. One possible explanation is that digital systems introduce elements of

interaction and immediacy that are typically absent in static assessment formats. When learners engage with reading materials through digital interfaces, the experience often becomes more dynamic and structured, which may encourage sustained attention and participation. Research on digital learning environments has consistently highlighted the role of interactive technologies in promoting learner engagement and facilitating deeper processing of instructional content [15].

The analysis of gain scores provides further insight into the observed improvements. Students in the experimental group achieved considerably larger gains compared with those in the control group. Because both groups were exposed to the same reading content and assessment format, the improvement cannot simply be attributed to familiarity with the test items. Rather, the results suggest that the structured digital environment of the PEER platform played an important role in supporting learning during the assessment process. Digital assessment tools often allow educators to organize tasks more effectively and track student progress in real time. Such features have been associated with improved formative assessment practices, as they enable teachers to monitor student performance more closely and make timely instructional adjustments [16]; [17].

Another notable finding relates to the improvement in non-linear reading proficiency. Students in the experimental group demonstrated slightly higher gains in this area compared with linear reading. This result is particularly relevant in light of contemporary literacy practices, where readers frequently encounter information presented in non-sequential formats such as charts, tables, infographics, and hyperlinks. Digital platforms naturally support these types of information structures, which may help learners develop the skills needed to interpret and synthesize complex sources. Recent studies on digital literacy emphasize that reading in online environments requires additional competencies beyond traditional comprehension, including the ability to locate, evaluate, and integrate information from multiple digital formats [18]; [19]. The PEER platform may therefore provide an assessment context that more closely mirrors the reading environments students encounter outside the classroom.

Beyond learning outcomes, the evaluation of the PEER platform revealed consistently high ratings across several software quality dimensions. Respondents strongly agreed that both the web and mobile applications demonstrated functional suitability, usability, security, and maintainability. These results are important because the success of educational technologies often depends not only on their instructional effectiveness but also on their usability and reliability in real classroom settings. Systems that are perceived as intuitive and easy to navigate are more likely to be adopted and sustained by teachers and students [20].

The positive evaluation of system security also deserves attention. In digital learning environments, the protection of student data is a critical concern for schools and administrators. The favorable ratings suggest that users trusted the platform's mechanisms for managing and storing information. At the same time, the strong perception of maintainability indicates that the system is flexible enough to accommodate updates and improvements over time. This adaptability is an important characteristic of educational software, as technological tools must evolve alongside changes in curriculum, pedagogy, and institutional needs [21]. Recent discussions in educational technology research further emphasize that scalable and adaptable digital systems are essential for supporting long-term integration in schools [22].

Taken together, the findings of this study highlight the potential value of digital assessment platforms in improving reading proficiency and enhancing the assessment experience. The significant gains observed among students who used the PEER platform, combined with the positive user perceptions of its quality and usability, suggest that such systems can serve as effective complements to traditional assessment methods. Rather than replacing conventional approaches entirely, digital platforms may offer additional opportunities for monitoring student progress, providing more engaging assessment experiences, and supporting data-informed decision-making in the classroom.

As schools continue to integrate technology into teaching and learning, the role of digital assessment tools is likely to become increasingly important. Platforms such as PEER demonstrate how technology can be used not only to measure learning outcomes but also to create assessment environments that align more closely with the digital contexts in which students read and interact with information. With thoughtful implementation, systems like PEER may contribute to more responsive and effective reading proficiency programs in secondary education.

VI. Conclusion

This study highlights how thoughtfully designed digital platforms can meaningfully improve the way reading proficiency is assessed in school settings. By streamlining assessment procedures and organizing results in a systematic manner, digital tools help schools establish more reliable and efficient reading comprehension assessment practices. More importantly, they allow teachers to spend less time on repetitive, manual tasks and more time interpreting results and responding to learners' needs. In this way, digital assessment platforms do not simply modernize existing practices; they strengthen the instructional support system that underpins reading development.

At a broader level, the findings suggest that technology-supported assessment is well aligned with the changing nature of reading literacy. As students interact with a wider range of texts, spanning linear, non-linear, and multimodal formats, digital platforms provide a flexible yet standardized approach to evaluating reading skills. The favorable evaluation of software quality across platforms also implies that when digital assessment tools are built in accordance with established quality standards, they are more likely to be accepted by users and sustained over time within educational institutions.

Looking ahead, future studies may build on this work by investigating how continued use of digital reading assessment platforms influences literacy growth over longer periods and across different grade levels. Expanding the scope of participants and implementing the system in multiple school contexts could further strengthen the applicability of the results.

There is also room to explore more advanced features, such as adaptive assessments, learning analytics, or AI-driven feedback mechanisms, to support more personalized reading interventions. Finally, qualitative inquiries into teachers' classroom practices and students' learning experiences would offer valuable insights into how assessment data from digital platforms are translated into concrete instructional improvements and meaningful learning outcomes.

ACKNOWLEDGMENT

The author sincerely thanks the teachers and students of Delfin Albano High School, under the leadership of Dr. Orlyne T. Demerin, for their support and valuable inputs. Appreciation is extended to the project team, namely Revilyn Totto, Christine Rondero, Vanessa Pagauitan, and Julius Lemu, for sharing PEER and their study on the developed platform.

REFERENCES

[1] N. B. Acedillo, "Improving the reading comprehension skills of Grade 5 pupils through contextualized learning materials: A school-based research," International Journal of Scientific and Research Publications, vol. 13, no. 3, pp. 380–392, Mar. 2023, doi: 10.29322/IJSRP.13.03.2023.p13545.

[2] OECD, "Programme for International Student Assessment (PISA)," Organisation for Economic Co-operation and Development. [Online]. Available: https://www.oecd.org/en/about/programmes/pisa.html

[3] J. Hu and Y. Wang, "Influence of students' perceptions of instruction quality on their digital reading performance in 29 OECD countries: A multilevel analysis," Computers & Education, vol. 189, p. 104591, 2022, doi: 10.1016/j.compedu.2022.104591.

[4] C.-M. Chen et al., "A web-based collaborative reading annotation system with gamification mechanisms to improve reading performance," Computers & Education, vol. 144, p. 103697, 2020.

[5] P. Auphan, A. Magnan, and J. Ecalle, "The high potential of computer-based reading assessment," Canadian Journal of Learning and Technology, vol. 46, no. 3, 2020.

[6] Chen, Danlei, and Gale Macleod. 2021. "Effectiveness of Digital Tools to Support Pupils' Reading in Secondary School: A Systematised Review." International Journal of Mobile and Blended Learning 13, no. 2 (January): 1–16. https://doi.org/10.4018/IJMBL.2021040101

[7] P. Delgado, C. Vargas, R. Ackerman, and L. Salmerón, "Don't throw away your printed books: A meta-analysis on the effects of reading media on reading comprehension," Educational Research Review, vol. 25, pp. 23–38, 2018.

[8] P. M. Stecker, L. S. Fuchs, and D. Fuchs, "Using curriculum-based measurement to improve student achievement: Review of research," Psychology in the Schools, vol. 42, no. 8, pp. 795–819, 2005, doi: 10.1002/pits.20113.

[9] N. Förster and E. Souvignier, "Learning progress assessment and goal setting: Effects on reading achievement, reading motivation, and reading self-concept," Learning and Instruction, vol. 32, pp. 91–100, 2014, doi: 10.1016/j.learninstruc.2014.02.002.

[10] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," MIS Quarterly, vol. 13, no. 3, pp. 319–340, 1989, doi: 10.2307/249008.

[11] ISO/IEC, ISO/IEC 25010:2011 Systems and software engineering—Systems and software Quality Requirements and Evaluation (SQuaRE)—System and software quality models. Geneva, Switzerland: International Organization for Standardization, 2011.

[12] Altamura, Lidia. 2024. "Turning the Page: What Research Indicates About Print vs. Digital Reading." Research Review 2, no. 4 (November 4, 2024 CDT).

[13] Totto, R., Rondero, R., Pagauitan, V., Lemu, J., and Tarun, I. 2024. PEER: A Web and Mobile Platform for Efficient Evaluation of Reading Proficiency. Unpublished capstone project.

[14] P. Black and D. Wiliam, "Assessment and classroom learning," Assessment in Education: Principles, Policy & Practice, vol. 5, no. 1, pp. 7–74, 1998.

[15] R. C. Clark and R. E. Mayer, E-learning and the Science of Instruction: Proven Guidelines for Consumers and Designers of Multimedia Learning, 4th ed. Hoboken, NJ, USA: Wiley, 2016.

[16] D. J. Nicol and D. Macfarlane-Dick, "Formative assessment and self-regulated learning: A model and seven principles of good feedback practice," Studies in Higher Education, vol. 31, no. 2, pp. 199–218, 2006.

[17] C. Redecker and Ø. Johannessen, "Changing assessment—Towards a new assessment paradigm using ICT," European Journal of Education, vol. 56, no. 1, pp. 79–96, 2021.

[18] D. J. Leu, E. Forzani, C. Rhoads, C. Maykel, C. Kennedy, and N. Timbrell, "The new literacies of online research and comprehension," Educational Psychologist, vol. 50, no. 1, pp. 37–59, 2015.

[19] P. Delgado, C. Vargas, R. Ackerman, and L. Salmerón, "Don't throw away your printed books: A meta-analysis on the effects of reading media on comprehension," Educational Research Review, vol. 35, p. 100433, 2022.

[20] J. Nielsen, Usability Engineering. San Francisco, CA, USA: Morgan Kaufmann, 2012.

[21] R. S. Pressman and B. R. Maxim, Software Engineering: A Practitioner's Approach, 9th ed. New York, NY, USA: McGraw-Hill, 2020.

[22] M. Bond, K. Buntins, S. Bedenlier, O. Zawacki-Richter, and M. Kerres, "Mapping research in student engagement and educational technology in higher education: A systematic evidence map," International Journal of Educational Technology in Higher Education, vol. 20, no. 1, pp. 1–21, 2023.