Diversitas Journal
ISSN 2525-5215
Volume 8, Número 4 (oct./dec. 2023) p. 2902 – 2921
https://diversitasjournal.com.br/diversitas_journal

The Design and Development of a Reading Assessment Manager

O Design e Desenvolvimento de um Gerenciador de Avaliação de Leitura

CRUZ, Sara Jane[1]; MENESES, Julius [2]

[1] 0000-0002-2818-0758; Rizal Technological Univaersity. Mandaluyong, NCR, Philippines. sjscruz@rtu.edu.ph.
[2] 0000-0002-0076-4962; Rizal Technological Univaersity. Mandaluyong, NCR, Philippines. jlmeneses@rtu.edu.ph.

A B S T R A C T

An alarming result from the 2018 Programme for International Pupil Assessment (PISA) revealed that the Philippines ranked lowest among the other 79 countries that participated in the assessment. This pushed the Department of Education to create reinforcement programs for reading, which were unfortunately shelved when face-to-face classes were canceled due to the COVID-19 pandemic. Hence, the researchers created the Reading Assessment Manager (RAM) for the continuity of reading programs such as the assessment of pupils' reading grade levels even during remote learning. In this study, RAM was created to virtually assess pupils' reading grade levels, recommend level-matching reading materials, and store data in individualized accounts. Following Type II Developmental Research Design, RAM was designed and developed using a series of needs assessment, prototyping, and user satisfaction of the system. Results of the administered System Usability Scale (SUS) showed that RAM outperformed 90-95% of industry software systems, while the Post-Study Usability Questionnaire (PSSUQ) reported an Above Average user satisfaction from the respondents.

R E S U M O

Um resultado alarmante do Programa de Avaliação Internacional de Alunos (PISA) de 2018 revelou que as Filipinas ficaram em último lugar entre os outros 79 países que participaram da avaliação. Isso levou o Departamento de Educação a criar programas de reforço para leitura, que infelizmente foram arquivados quando as aulas presenciais foram canceladas devido à pandemia de COVID-19. Portanto, os pesquisadores criaram o Gerenciador de Avaliação de Leitura (RAM) para a continuidade de programas de leitura, como a avaliação dos níveis de leitura dos alunos, mesmo durante o aprendizado remoto. Neste estudo, o RAM foi criado para avaliar virtualmente os níveis de leitura dos alunos, recomendar materiais de leitura compatíveis com o nível e armazenar dados em contas individualizadas. Seguindo o Design de Pesquisa de Desenvolvimento Tipo II, o RAM foi projetado e desenvolvido usando uma série de avaliações de necessidades, prototipagem e satisfação do usuário do sistema. Os resultados da Escala de Usabilidade do Sistema (SUS) administrada mostraram que o RAM superou 90-95% dos sistemas de software da indústria, enquanto o Questionário de Usabilidade Pós-Estudo (PSSUQ) relatou uma satisfação do usuário acima da média por parte dos respondentes.

ARTICLE INFORMATION

Histórico do Artigo:
Submetido: 28/09/2023
Aprovado: 28/11/2023
Publicação: 08/12/2023

Keywords:
Reading Assessment Software, Reading Comprehension, Educational Technology

Palavras-chave:
Avaliação de Leitura, Compreensão de Leitura, Tecnologia Educacional

DOI: 10.48017/dj.v8i4.2790

Introduction

Over the past decades, more and more people are crossing over the illiteracy line and moving upward the literacy stage. Roser and Ospina (2018) stressed from a historical viewpoint that the world population literacy levels have significantly risen in a couple of centuries. For 65 years, the global literacy rate continues to improve by 4% every five years—from 42% in 1960 to 86% in 2015. But underlying these success rates in educating the global population lies a research gap that poses the question, how extensively do one's reading comprehension skills affect and improve one's literacy?

Baier (2005) stressed that reading comprehension is an essential foundation skill of literacy that is a critical indicator of educational success. In turn, Keyser (2018) argued that reading comprehension is a crucial link to effective reading that is essential for a rich academic, personal, and professional life. This highlights the importance of reading comprehension in every classroom as a must for every academic institution.

In the 2018 Programme for International Student Assessment (PISA 2018), student participants in the Philippines scored beneath the standard rating set by the Organization for Economic Cooperation and Development (OECD). It left the country ranking lowest among the other 79 countries.

The result pushed the Department of Education to enhance and strengthen the implementation of reading programs such as the Philippine Informal Reading Inventory (PhilIRI). According to Llego (2019), Phil-IRI is a reading assessment instrument for identifying the reading grade levels of the learners in terms of oral reading, silent reading, and listening comprehension. The assessment has two variations, one for assessing reading comprehension in English and the other in Filipino. The assessment results aim to guide the teachers in conceptualizing personalized reading programs per student.

Since the Phil-IRI is an assessment that requires face-to-face interaction with the learner, setbacks challenged the test administration, especially during the pandemic when classes migrated to virtual instruction. Teachers found the exam impossible to be administered online and tedious for both the pupil and the test administrator. Hence, schools shelved reading programs utilizing Phil-IRI for the time being.

With the emergence of the 4th Industrial Revolution and the cyber-physical system of things, such dilemma can be mitigated using existing technologies. Using current technologies, the traditional pen-and-paper reading assessments can be transferred online, making them virtually accessible for learners, even in the comforts of their homes. It is from this standpoint, that the researchers developed RAM for Basic Education.

RAM is a web-based learning management system that can virtually store a reading inventory of learners' reading grade levels and suggests level-matching reading materials by mining the system's database. Results given by RAM can significantly help academic institutions in constructing their individualized reading programs.

Methodology

The Developmental Research Design was utilized by the researchers in the creation and evaluation of the Reading Assessment Manager as it is a "systematic study of designing, developing, and evaluating instructional programs, processes, and products that must meet criteria of internal consistency and effectiveness" (Richey & Nelson, 2001). Since the emphasis of the study is on the generation of context-specific conclusions through undertaking the product design, development, and evaluation of analyzing the conditions that facilitate the use of a specific product (Ibrahim, 2016), the researchers followed Type 1 Developmental Research Design. Adding to this, Akker (1999) states that Type 1 Developmental Research Design can be complemented with an instructional system design that is often cyclical wherein analysis, design, evaluation, and revision activities were iterated until the prototype brings about a balance between the ideals and the realizations. Thus, Figure 1 below shows the contemporary instructional system design model called Rapid Prototyping with the phases stressed by Akker (1999).

Figure 1.
Schematic Paradigm of Rapid Prototyping.

Source: (Gutierrez, 2015).

According to Gutierrez (2015), Rapid Prototyping is a non-linear instructional design of design, development, and evaluation. It is rooted in the typical Analysis, Design, Develop, Implement, and Evaluate (ADDIE) model that only provides feedback at the end of the instructional design— thus making the process painstaking in certain situations where revisions or corrections in the preliminary parts of the system are needed.

Figure 2.
Conceptual Framework of the Research

Source: (Gutierrez, 2015).

In this research, the study combined the principles of Rapid Prototyping and the phases of Type 1 Developmental Research Design according to Akker (as cited by Ibrahim, 2016), which includes analysis, development, formative evaluation, summative evaluation, and reflection on the development methodology. It would support the whole research process and evaluation, as reflected in Figure 2.

Participants of the Study

Participants for the Needs Assessment

Using purposive sampling, the researchers only gathered data from the concerned school personnel, namely the five Language Arts teachers from the research locale who would use the system and their three coordinators (Lower Grade School, Upper-Grade School, SPED) who would lobby the implementation of the Reading Assessment Manager. This group represents the client of the system and the members of the Focus Group Discussion.

Respondents for the Usability Testing Using SUS and PSSUQ The researchers utilized a "problem discovery sample size formula" (Vizri, as cited by Lewis, 1994). This sample size formula is usually used by software developers for usability testing, according to Vizri. It claims that: (1) observing four (4) or five (5) participants will allow a usability practitioner to discover 80% of usability problems of a product, (2) observing additional participants will reveal fewer and fewer new usability problems, and (3) more severe usability problems are easier to detect with the first few participants." Therefore, using many participants is not ideal in system usability evaluations.

Using the problem discovery sample formula to get an appropriate number of respondents, results revealed that 22 respondents were needed for the usability testing. This group was composed of 12 student respondents from the intermediate level, five Language Teachers, three School Coordinators, and two Software Engineers. They evaluated the produced Reading Assessment Manager in terms of usability, accessibility, and appearance. From the gathered data, a test of difference was done to know if there is a significant difference between the manual test administration and interpretation vis-à-vis the computer-aided one in terms of resource allocation and needed user skills for the test administration and interpretation as perceived by the teachers.

These 22 respondents evaluated the developed system by answering two questionnaires such as the System Usability Scale (SUS) and the Post-Study System Usability Questionnaire (PSSUQ).

Research Instruments

Phase 1: Analyze

In this phase, the researchers aimed to get from the teacher-respondents the reading comprehension needs of the pupils. It was brought out using the following instruments:

Front-End Analysis through Focus Group Discussion

This focus group discussion aimed to know the high-impact performance problem, identify specific performance deficiencies, identify the root cause of the performance problems, and increase the chances for the selection of appropriate and cost-effective solutions.

In this discussion, the researchers asked the 5 Language Arts Teachers and 3 School Coordinators analysis questions using Harless' 13 Smart Questions for Front-End Analysis (FEA). This FEA is broken into the following 3 phases for a more thorough questioning:

1. Performance Analysis- aimed at isolating and identifying performance problem
2. Cause Analysis- aimed at identifying and isolating the underlying reasons for the performance problem
3. Intervention Selection, Design, and Implementation- aimed at determining a cost-effective solution to address, eliminate, minimize, reverse, or prevent the root causes of the performance problem

Needs Analysis Survey

The needs analysis survey administered brought about the focused purpose of the proposed system and its needed system requirements in terms of assessment content, data banking, and data presentation. A Google Form of this interview survey was sent to the 5 Language Arts teachers and 3 School Coordinators. For the validation of this qualitative tool and its results, member checking was done with the respondents.

Instructional Goal Setting

To solidify the goals and objectives of the Reading Assessment Manager, an Instructional Goal Setting Interview Questionnaire was given to the 3 School Coordinators. This questionnaire aimed to finalize the goals and expectations for the Reading Assessment Manager to be developed. Based on the given Front-End Analysis and Needs Assessment Survey, the cluster coordinators chose the objectives and goals adopted.

For validation of this tool, member checking was also conducted before the respondents affix their signatures on the final report.

Phase 2: Design

In this phase, the researchers used the results collated from the tests administered using the instruments shown in phase 1.

Phase 3: Develop and Produce

In this phase, the researchers used the Phil-IRI parameters to get the reading grade level of the grades 2-6 pupils.

Phase 4: Evaluate

In this phase, the researchers aimed to get from the respondents the evaluation ratings and feedback on the Reading Assessment Manager. Data were gathered using the following instruments:

System Usability Scale

System Usability Scale (SUS) is a standardized questionnaire developed by John Brooke in 1986 and was improved (version 3) in 1996. This instrument is a public domain tool with a Cronbach Alpha Level of 0.92, according to Sauro (2018). For this study, this instrument was administered to the 5 Language Arts Teachers, Coordinators, and 12 Pupil Participants.

5. Appeal and Accessibility Questionnaire

An Appeal and Accessibility Interview Questionnaire was given to the 5 Language Arts teachers and 3 School Coordinator-respondents. On the other hand, a semi-structured interview of the 12 pupil-participants, guided by the same questions from the questionnaire, was also conducted. Member checking was done to validate the results of the questionnaire.

6. Post-Study System Usability Questionnaire (PSSUQ)

This survey aims to measure users' perceived system satisfaction in terms of System Quality, Information Quality, and Interface Quality. This was given to the Language Arts Teachers, Cluster Coordinators, Software Engineers, and Pupil participants. The PSSUQ is a standardized questionnaire developed by Jim Lewis in the 1980s.

7. Reading Test Administration Survey

This survey aims to examine the significant difference between the traditional administration of a reading assessment against the automated as perceived by the teacher-respondents. This researcher-made survey was validated by three experts in the field who were asked to rate each question item whether it is "essential;" "useful but not essential;" or "not essential." From an initial 20-item survey, the list was narrowed down to 10 questions which showed acceptable levels of significance after getting each question item's Content Validity Ratio (CVR).

Furthermore, the instrument was tested for its reliability by getting the Cronbach Alpha of the survey after conducting pilot testing with five teachers of the other grade levels. This group of teachers, as chosen, were not from the sample participants of the research. This is to avoid the contamination of the research sample data gathered (Peat, et.al, 2002).

After conducting the pilot testing, the results showed a Cronbach Alpha of 1.00, or with excellent internal consistency, according to Cronbach's (1951) Internal Consistency Values.

Results

Objective and Scope of the Reading Assessment Manager According to the Focus Group Discussion Administered

The following data show the scope and objectives of RAM based on the responses of the participants in the conducted Front-End Analysis Interview, Instructional Goal Setting Questionnaire, and Needs Analysis Survey.

a. Target Skill Outcomes

Table 2.

Target Skill Outcomes

Skills	No. of Responses	Percentage
Good reading speed	0	0%
Good oral reading	2	25%
Good reading comprehension	6	75%

Source: Own authorship.

Based on the Needs Assessment survey, as shown in Table 2, 75% of the language teacher and coordinator respondents prefer to develop pupils' reading comprehension. According to one of the language teachers-respondents, aiming to develop pupils' reading comprehension helps students understand the text read.

Subsequently, this claim was strengthened by the conducted Front-End Analysis Interview with the Coordinators and Language Arts Teachers. The interview revealed that most of the students struggle to comprehend texts and selections in Filipino as perceived by 50% of the interview participants.

b. Finalized Goals and Objectives

Given the set of problems at hand and the result of the survey which mainly revealed the respondents' aim to focus on reading comprehension, the coordinators finalized the scope and objectives of the Reading Assessment Manager in an Instructional Goal Setting Questionnaire given by the researchers. The following are the set of goals and objectives from the coordinator-respondents:

(1) Goal: Develop pupils who have a deep understanding of Filipino selections

The system was designed to assess pupils' understanding of the reading texts and selections in Filipino. By designing the system to showcase and gather reading comprehension data of the grades 2-6 pupils, teachers would be able to identify pupils' level of understanding of what was read. This would give the Language Arts teachers baseline data for any intervention that they would aim to apply.

(2) Objective 1: To be able to develop pupils' reading comprehension

The main purpose and feature of the Reading Assessment Manager are to help the teachers develop pupils' reading comprehension through reading test administration and data banking for plotting each learner's reading progress.

Objective 2: To be able to have reading skills appropriate to their level

The Reading Assessment Manager as designed, aimed to gather data that can reveal pupils' reading levels. By doing so, the teachers would have baseline data on the students reading grade level which they can use to determine which technique or method is appropriate as an intervention. From here, the Language Arts teachers would also know if the pupils were at par with their expected reading grade level and if not, they would be able to identify students who need remediation.

Objective 3: To be able to understand the meaning of the text read

Administration and collection of results in the reading assessment tests as the main feature of the Reading Assessment Manager, would check and validate pupils' progress in understanding the text read.

System Requirements According to the Needs Analysis Survey and Instructional Goal Setting Administered

According to the Needs Assessment Survey and Instructional Goal Setting Questionnaire given to the Coordinators and Language Teachers of Grades 2 to 6 pupils, the gathered information yielded the given results as shown in Table 13 below.

a. On Data to be Gathered

Table 3.

Data to be Gathered

Data	No. of Responses	Percentage
Pupil's reading speed	0	0%
Pupil's oral miscues	3	37.5%
Pupil's reading comprehension level	5	62.5%

Source: Own authorship.

Table 3 shows that 62.5% of the respondents prefer to gather data on pupils' reading comprehension levels. According to one of the respondents, this data would help the teachers identify which skills to develop to help the students. Adding to this, another respondent claimed that this data would guide the teachers in tailor fitting students' activities.

On the other hand, 37.5% of the respondents preferred to gather pupil's oral miscues while no respondent (0%) chose pupil's reading speed. With these results, the Reading Assessment Manager was designed to assess the pupil's reading comprehension level.

b. On Frequency of Test Administration

Table 4.

Frequency of test administration

Frequency	No. of Responses	Percentage
Once a year	0	0%
Every quarter	5	62.5%
Once a month	3	37.5%

Source: Own authorship.

Table 14 illustrates that 62.5% of the respondents preferred to conduct the assessment every quarter, while 37.5% preferred to conduct the aforementioned once a month. According to one of the respondents who chose the "every quarter" test administration, this timeframe is just accurate to track the skills development of the students. Another respondent who answered the same agreed that this time span would be long enough to see if the techniques or interventions to be given are working.

Guided by the result of the survey, the Reading Assessment Manager was designed to gather test scores every quarter as preferred by the majority of the respondents.

c. On Data Banking

Table 5.

Data Banking

Type of Data	No. of Responses	Percentage
Individual Data Progress	1	12.5%
Class Data Progress	0	0%
Grade Level Data Progress	1	12.5%
All of the Above	7	87.5%

Source: Own authorship.

Table 5 shows that most of the respondents, with a percentage of 87.5%, opted to know Grades 2-6 pupils' Individual Data Progress, Class Data Progress, and Grade Level Data Progress. On the other hand, 12.5% chose to only know the Pupil's Individual Data Progress and 12.5% also chose to only know the Grade Level Data Progress of the pupils in reading and comprehension.

By getting all the criteria data, the teachers and administrators would be able to have a comprehensive view on the pupils' reading and comprehension level. This in turn, could help the school in conducting reading programs and other interventions.

d. On Data Presentation

Table 6.

Data Presentation

Type of Graph	No. of Responses	Percentage
Data in bar graph	5	62.5%
Data in line graph	3	37.5%
Data in pie graph	0	0%

Source: Own authorship.

For a better understanding of the data gathered, the respondents were asked in the Needs Assessment Survey, how they expect the Reading Assessment Manager to present the assessment results. Based on Table 16, 62.5% preferred to view the results as presented through a bar graph. While 37.5% and 0% chose that the data should be presented using a line graph and pie graph, respectively. With this result, the Reading Assessment Manager created was designed to present data through a bar graph.

e. Additional System Requirements

To fully tailor-fit the Reading Assessment Manager, the coordinators were also asked for additional system requirements that they expect the system to have. The coordinators expected that the system would (1) include pictures in the system for the pupils to enjoy the whole experience; (2) make the stories interesting; (3) recommend grade-level-appropriate reading materials for the pupils to be able to connect with the reading selections and texts; (4) provide feedback section that students can check to monitor their progress; (5) generate a chart or table to show pupils' progress; and (6) design the system that can be used offline.

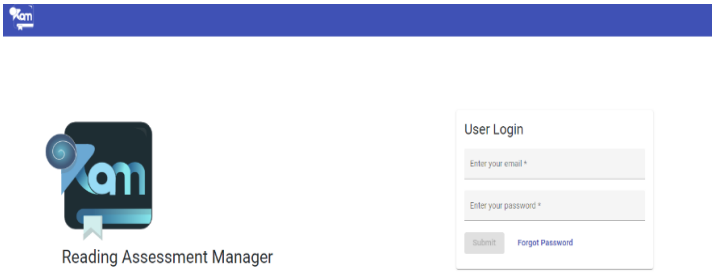
Features of the Reading Assessment Manager According to the Needs Assessment and the Additional System Requirements

As requested by the respondents from the focus group discussion, the following features were integrated in the Reading Assessment Manager:

Both the student account and the teacher's account start with the log-in page. Since one of the goals of the system is to generate a report and feedbacking mechanism that is based on the result of each pupil, individualized user accounts were given to the students. Likewise, teachers were also given user login details, for them to have access on the data of their handled class. Each student and teacher can access RAM using assigned user log-in details.

Figure 2.

Log-in Page a.1 Objective 1: Develop pupils' understanding of Filipino selections



Source: Personal archive.

Reading Assessments for both English and Filipino were integrated into the system to meet the requirements of the client which focuses on developing a deeper understanding of Filipino selections.

Figure 3.

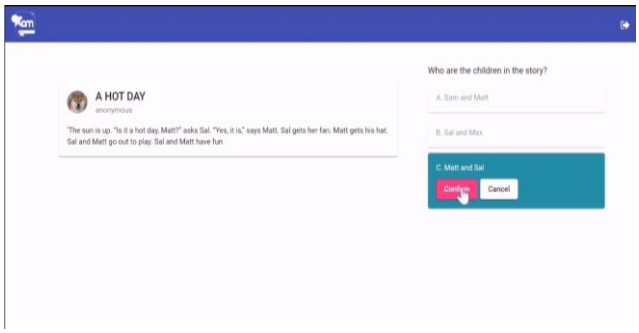
Test Selection Page a. 2. Target Skill Outcomes: Good reading comprehension

Source: Personal archive.

A reading and comprehension test instrument designed to reveal the test takers' reading grade level was used in the system as its main content.

Figure 4.

Reading Selection with Question Item a.3. Objective 2: Identify pupils' reading grade level

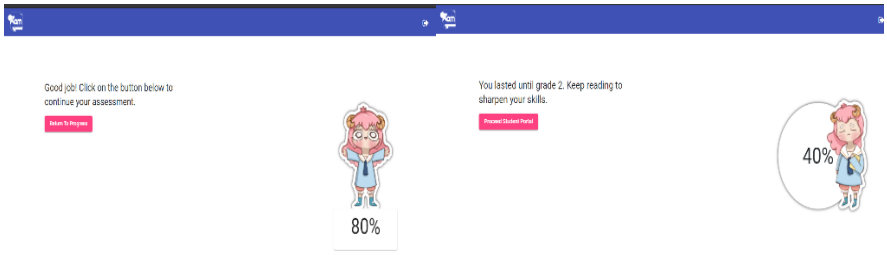


Source: Personal archive.

The system algorithm was designed based on the formula indicated in the Phil-IRI. Ideally, the pupils start taking the test assigned for their current grade level and continue to the next higher or lower level until they reach their "Frustration Level," equivalent to scoring below 80%. A feedback page is designed to guide the test takers on the next step to take.

Figure 5.

Feedback Page a.4. System Requirements: On Data to be Gathered



Source: Personal archive.

According to the clients, the system should be able to show the reading comprehension level of the test takers. This feature was added to the system under the "Student Portal Page" for the students and the "Summary Page" for the teachers.

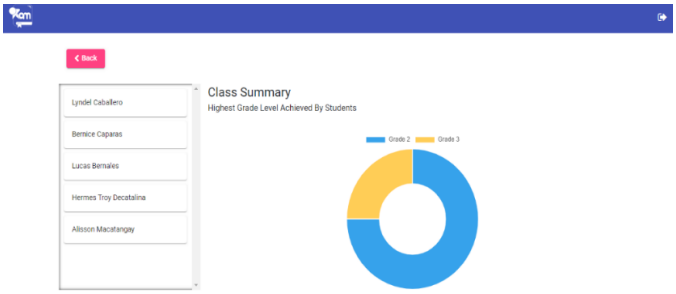
On the Student Portal Page, data can be seen in bar graph form. Further information on the test takers' scores can be accessed by hovering the mouse pointer above the bar graph. The pupil's data is aggregated according to language, quarter, set, and school year as requested by the client. This feature also satisfies the additional requirement of the clients, which is to provide a feedback section that students can check to monitor their progress.

Figure 6.

On the other hand, the teachers can access the generated reports of the system under their accounts. Using their own user details, they can have personalized access to the data of the pupils that they are teaching. This can be seen in their "Class Summary Page" and "Individual Student Summary Page." This feature satisfies the requirement of the clients, which is to generate a chart or table to show pupils' progress.

Figure 7.

Class Summary Page under the Teacher Account



Source: Personal archive.

Figure 8.

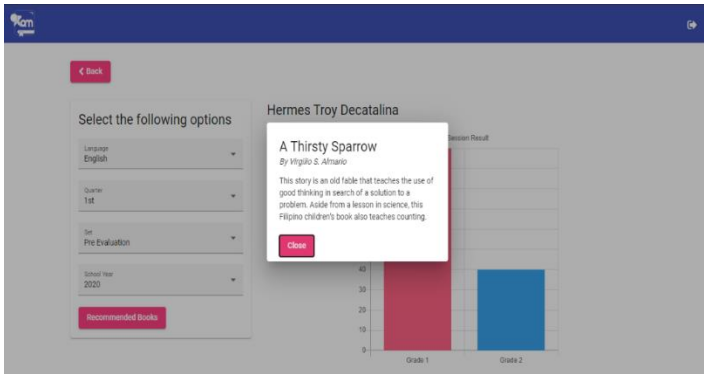
Individual Student Summary Page under the Teacher Account

Source: Personal archive.

Additional Feature of the System: Recommend grade level appropriate reading materials for the pupils to be able to connect with the reading selections and texts.

Figure 9.

Recommended Reading Materials



Source: Personal archive.

The researchers tied up with Adarna Publishing for the set of book recommendations appropriate for the pupils' reading grade levels. This can be accessed by the test takers under their "Student Portal Page." The same feature can be seen in the Teacher's Account, in the "Individual Student Summary Page."

Discussion

System Usability Scale (SUS) Results

Aiming to examine the ease of use in terms of learnability and operational efficiency through system interaction, the researchers administered the System Usability Scale (SUS). Evaluated by 22 respondents, the test garnered the following results.

Table 7.

System Usability Scale (SUS) results

Respondents	Score	Interpretation
Experts	93.8	Best Imaginable
Teachers and Coordinators	81	Excellent Usability
Pupils	79.4	Excellent Usability
Over-all	81.6	Excellent Usability

Source: Own authorship.

Table 7 reflects that among the experts, the usability of the system compared to the other web-app systems can be considered as Best Imaginable (93.8). On the other hand, for the Language Arts Teachers, Coordinators, and Pupils, it has Excellent Usability with SUS Scores of 81 and 79.4. The overall results uncover that RAM received a score of 81.6 which equates to excellent usability and a Grade of A or superior performance (Sauro, 2018). Referring to the SUS curved grading scale (Sauro, 2011), a SUS Score of 81.6 also equates to 90-95 percentile.

Table 8.

SUS Curved grading scale

SUS Score Range	Grade	Percentile Range
0-51.7	F	0-14
62.7-64.9	D	15-34
65-71	C-	35-40
71.1-72.5	C+	60-64
72.6-74	B-	65-69
74.1-77.1	B	70-79
77.2-78.8	B+	80-84
78.9-80.7	A-	85-89
80.8-84	A	90-95
84.1-100	A+	96-100

Source: Sauro (2011).

The Curved Grading Scale interpretation of SUS scores is based on the data uncovered from more than 10,000 respondents, hundreds of collected usage data for 30 years, and hundreds of products such as hardware, software apps, websites, mobile apps, or voice user interfaces (Brooke as cited by Sauro, 2018). Therefore, the percentile range of 90-95 of RAM indicates that the system scores 90-95% better than other systems in the database.

Points for improvement in terms of appeal, and accessibility

a. Overall look of the System

Respondents have conveyed that the color scheme is appealing and suitable for young learners. Likewise, the font style is readable, but the font size is small for half of the respondents and just fine and readable for the other half.

b. Recommendations for the improvement of the overall appeal

20% of the respondents expressed that it would be better if there would be more pictures, especially for the reading passages. This is to help the students get more engaged with the reading assessment. Also, in line with the illustrations, 1 respondent commented on the choice of the system mascot. She states that some students are sensitive when it comes to fictional drawings. It would be safer if the system would showcase non-fictional characters.

Also, one respondent commented on the "Log Out" sign. She recommends making it more visible.

c. Best feature of the system's look

According to the respondents, the system looks clean and organized. It uses child-friendly illustrations with calm colors. The look is simple and showcases minimal distractions.

d. Graphics and design

Three (3) out of the 10 respondents conveys that the graphics and design used in the Reading Assessment Manager look good and the rest stated that it looks appealing and motivating. One (1) of the respondents also added that the illustrations help in engaging the pupils with the system.

e. Accessibility in terms of logging-in

The respondents commented that logging in was easy for the users since they just need to input their school email address and password. It was simple and easy to remember.

f. Access to reports

According to the participants, accessing and generating their class' and students' reports are easy and simple. As stated by one of the respondents, navigating was easy for there are fill-out guide sections for what data they would like to generate.

g. Recommendations for accessibility improvements

Most of the respondents do not have any more recommendations since this web-based system application can be accessed online and offline. According to one of the teacher-respondents, it would be better if this would be used offline and just installed in the school computer lab so the teachers can be with the pupils as they take the exam. This would give more authenticity to the child's responses and test scores since the teachers know that it is the

actual learners who answered the test, as opposed to giving it to the students to be taken at home wherein the teachers would not know the authenticity of the test results.

h. Accessibility problems encountered

All the respondents accessed the Reading Assessment Manager using their laptops. According to them, no major problems were encountered except for one respondent who entered the wrong email and password. She stated that a log-in error notification appeared, and it helped her in re-entering her details.

i. Recommendations for the problem encountered.

Since there were no major problems encountered while accessing the system, the respondents stated that the Reading Assessment Manager is already fine as it is.

j. Comments on the overall accessibility

The respondents expressed their satisfaction with the created Reading Assessment Manager since it can be accessed online and offline. On the other hand, some respondents just commented that the system should also be accessible through mobile and tablet.

Usability results as revealed by the PSSUQ

Using Post-Study System Usability Questionnaire (PSSUQ), RAM was also evaluated by the respondents after the beta-testing and system walkthrough. Similar to SUS, PSSUQ is a 16-item non-functionality usability questionnaire, developed by the IBM Design Center in 1992 and designed as a scenario-based usability test for a more targeted evaluation (Rotolo, 2017). It examines the system's usefulness, information quality, and interface quality. PSSUQ was administered in this study, alongside SUS, for a more thorough insight into the usability of RAM. The tool garnered evaluation results from the participants as shown in the table below.

Table 9.

Post-Study System Usability Questionnaire (PSSUQ) results

Respondents	Score	Interpretation
System Usefulness	2.65	Above Average
Information Quality	2.92	Above Average
Interface Quality	2.38	Above Average
Over-all	2.67	Above Average

Source: Own authorship.

According to Lewis (1993), the instrument results are read invertedly wherein the lower the mean value is, the better the usability rating. Thus, as seen in Table 9, RAM received an overall rating of Above Average (2.67). Per sub-scale, all garnered Above Average mean scores of 2.65 in System Usefulness, 2.92 in Information Quality, and 2.38 in Interface Quality. Among the three (3) sub-scales, Interface Quality garnered the highest rating. Conversely, Interface Quality has also garnered the lowest rating in its Question item number 7 which is

"The system gave error messages that clearly told me how to fix my problems." This item got a PSSUQ score of 4.29 which is Below Average. This means that more effort should be made in designing a good error message.

Utilizing the PSSUQ results, a regression analysis was undertaken to show which of the three (3) sub-scales affect the overall result of the respondents' evaluation.

Table 10.

Regression Analysis of the factors affecting the overall results of Post Study System Usability Questionnaire (PSSUQ)

Model		**Unstandardized Coefficients**		**Standardized Coefficients**	**t**	**Sig.**
		B	Std. Error	Beta		
1	(Constant)	0.00048	.003		.178	.860
	System Usefulness	.375	.000	.465	884.355	.000
	Information Quality	.375	.001	.432	744.390	.000
	Interface Quality	.250	.001	.197	300.149	.000
a. Dependent Variable: Overall Result of PSSUQ						
R-squared = 1.00						
F-Value = 3079765.022						
p- value= .000						
alpha =.05						

Source: Own authorship.

Multiple regression was conducted to predict which of the three (3) factors affected the overall result of the PSSUQ the most. The multiple regression rendered an overall R2 of 1.00 or 100% which is expected because the three (3) predictors are just portions of a whole which is the overall result. This means that no external or other factors will be accounted for except the three predictors namely: System Usefulness, Information Quality, and Interface Quality.

A significant regression equation was found (F (3079765.022), p =.000), with R2 of 1.00 or 100% of the effects in the overall results are attributed to the three (3) given predictors. Using the B coefficients, this equation of determination was formed which states that the Overall result = .000486+.375(Systems Usefulness) +.375(Information Quality) +.250(Interface Quality). This means that every one-point increase in Systems Usefulness corresponds to a .375 increase in the overall results which is the same as the Information Quality. In the interface Quality, the model tells us that every one-point increase in its score renders a .250 increase in the overall results.

All three (3) interventions or predictors have a significant effect on the overall results which is the dependent variable in this study as what the p (Sig.) values reported in the table. They are all less than the alpha level of .05 and reflected as identical .000 but that does not mean that they are equal. Of the three (3) predictor variables, System Usefulness and Information Quality are the two (2) factors affecting the most the overall results with an identical percentage effect of 37.50% leaving Interface Quality with only 25.00%.

Significance testing of Manual Test Administration vis-à-vis Automated Test administration of the Reading Assessment Manager

The main goal of this research is to develop a Reading Assessment Manager for Grades 2-6 Pupils of a Progressive School in Metro Manila, but aside from this, one of its purposes is for efficient human resources and budget allocation. Hence, the researcher also checked if as perceived by the teachers and school coordinators, this created Reading Assessment Manager helps in the efficient management of human and capital resources. To measure this, a researcher-made survey was given to the five (5) Language Arts teachers and three (3) School Coordinator respondents. Using paired sample t-test, the results revealed that the mean difference of 1.00 between the Manual Test Administration (2.65) and Automated Test Administration (1.65) of the Reading Assessment is statistically significant since the p-value (.010) is less than the specified α level (.05). Likewise, this means that there is enough evidence to reject the null hypothesis and accept its alternative hypothesis. This 1.00 reduction explains to us that Automated Administration outperformed the Manual Administration of the Reading Assessment.

Conclusion

Based on the given findings, the data revealed that the developed RAM performs 90-95% higher than other systems in the industry and was rated by the respondents with Above Average user satisfaction. Furthermore, Language Arts Teachers and School Administrators have perceived that using RAM is more efficient than implementing the reading assessment manually. Automation of the learning experiences such as what RAM did, as perceived by the respondents, can help school administrators in efficiently allocating human and capital resources. As a shared insight during the focus group discussion, the teachers have also highlighted that having such a Reading Assessment Manager that is available online and offline, can show the responsiveness of the school to the current scenario and needs of the pupils. Now, more than ever, school leaders need to strengthen their technology departments to employ flexible and innovative solutions to adapt to changing times.

Due to this, there is a positive likelihood that RAM can be seamlessly adopted in the research locale, but after some further improvements in the Interface Quality, specifically on the issuance of a good error message.

Limitations

This study is only limited to the design and development of the Reading Assessment Manager based on the system requirements and needs assessment of the target users in the research locale.

Ethics Statement

Participants of this study have provided their informed consent on the data gathering, data management, and data disposal procedure to be implemented in undertaking the research.

References

Akker, V. (1999). *Principles and methods of development research*. Retrieved http://www.heybradfords.com/FormativeResearchInstructionalUnit/Van%20der%20Akker%20Ch1.pdf

Alcober, N. (2020, February 11). *Pasig schools focus on reading*. Retrieved from https://tribune.net.ph/index.php/2020/02/11/pasig-schools-focus-on-reading/

Awabdy, G. W. (2012). *Background knowledge and its effect on standardized reading Comprehension Test Performance*. Retrieved from https://escholarship.org/content/qt7ph9r3ts/qt7ph9r3ts_noSplash_58f02b04b879be3cf2f9eb63de267952.pdf

Baier, R. J. (2005). *Reading Comprehension and Reading Strategies*. Retrieved from https://core.ac.uk/download/pdf/5066651.pdf

Biancarosa, G., & Griffiths, G. (2012). *Technology tools to support reading in the digital age*. Retrieved from https://files.eric.ed.gov/fulltext/EJ996196.pdf

Beins, Bernard C. (2014). *APA Style Simplified: Writing in Psychology, Education, Nursing, and Sociology (Wiley Custom Edition)*. John Wiley and Sons Inc.

Department of Education. (2018). *The Philippine informal reading inventory manual 2018*. Retrieved from http://www.davaocitydeped.ph/downloads/forms/313-phil-iri-full-package.html

Gutierrez, K. (2015, July 9). *The ins and outs of Rapid Prototyping for eLearning*. Retrieved from https://www.shiftelearning.com/blog/rapid-prototyping-for elearning#:~:text=Rapid%20Prototyping%20is%20an%20instructional

Ibrahim, A. A. (2016, January 10). *Definition Purpose and Procedure of Developmental Research: An Analytical Review*. ResearchGate. Retrieved September 30, 2021, from https://www.researchgate.net/publication/312934049_Definition_Purpose_and_Procedure_of_Developmental_Research_An_Analytical_Review.

Kenton, W. (2020, September 21). *How Multiple Linear Regression Works*. Retrieved December 06, 2020, from https://www.investopedia.com/terms/m/mlr.asp

Keyser, A. (2021). *Why is Reading Important*. Retrieved from https://www.worksheetcloud.com/blog/why-is-reading-important/

Lethaby, C. (2005, April 14). *What is linguistics knowledge?* Retrieved from https://rafangel.wordpress.com/2005/04/14/bella/

Lewis, J. (1994). *Sample Sizes for Usability Studies: Additional Considerations*. Retrieved December 05, 2020, from https://www.researchgate.net/publication/15128047_Sample_Sizes_for_Usability_Studies_Additional_Considerations

Lewis, J. (2002, September). (PDF) *Psychometric Evaluation of the PSSUQ Using Data from Five Years of Usability Studies*. Retrieved December 06, 2020, from https://www.researchgate.net/publication/220302199

Llego, M. (2019, March 4). *Revised Philippine informal reading inventory (Phil-IRI)*. Retrieved from https://www.teacherph.com/revised-phil-iri/

Mcginnis, D. (2018, December 20). *What is the fourth Industrial Revolution?* Retrieved from https://www.salesforce.com/blog/2018/12/what-is-the-fourth-industrial-revolution-4IR.html

McRae, S. (2012). *Utilizing the Interactive Reading Model in a Continuing Education Course*. Retrieved from https://digitalcollections.sit.edu/cgi/viewcontent.cgi?article=1541&context=ipp_collection

Moore, A. L. (n.d.). *A research review of cognitive skills, strategies, and interventions for reading comprehension*. Retrieved from http://download.learningrx.com/reading-comprehension-research-paper.pdf

Moran, K. (2019, December 1). *Usability Testing 101*. Retrieved December 05, 2020, from https://www.nngroup.com/articles/usability-testing-101/

Mustaro, P. N. (2007*). Structure of storyboard for interactive learning objects development*. Learning objects and instructional design, 253-280.

Organisation for Economic Co-operation and Development. (2018). *Philippines - OECD. OECD Programme for International Student Assessment (PISA)*. Retrieved April 19, 2022, from https://www.oecd.org/pisa/publications/PISA2018_CN_PHL.pdf

Peat, J., Mellis, C., Williams, K. and Xuan W. (2002), *Health Science Research: A Handbook of Quantitative Methods*, London: Sage.

PSSUQ (Post-Study System Usability Questionnaire). (2020, April 02). Retrieved December 06, 2020, from https://uiuxtrend.com/pssuq-post-study-system-usability-questionnaire/

Razi, o., & Grenfell, J. (2012). *The impact of linguistic knowledge on learner strategy deployment*. Procedia - Social and Behavioral Sciences, 47, 818-822.

Richey, R. (1994). *Developmental Research Design Research: The definition and scope*. Retrieved from https://eric.ed.gov/?id=ED373753

Richey, R. C., & Nelson, W. A. (2001, August 3*). Developmental Research. 42: Developmental research*. Retrieved April 19, 2022, from http://members.aect.org/edtech/ed1/42/index.html#:~:text=Developmental%20research%2C%20as%20opposed%20to,127).

Roser, M., & Ortiz-Ospina, E. (2023). *Literacy*. Retrieved from https://ourworldindata.org/literacy

Rotolo, T. (2017, February 24). SUS and PSSUQ: *Quantifying user experiences*. TryMyUI Blog. Retrieved April 19, 2022, from https://www.trymyui.com/blog/2017/02/24/sus-pssuq-website-usability-surveys/

Sauro, J. (2018, September 19). *5 Ways to Interpret a SUS Score*. Retrieved December 06, 2020, from https://measuringu.com/interpret-sus-score/

Snow, C. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*.Retrieved from https://www.rand.org/pubs/monograph_reports/

SOAR (Scholastic online reading Assessment. (n.d.). *The Best Choice to Measure Reading Comprehension and Progress*. Retrieved from https://education.scholastic.ca/category/SRI_READING_INVENTORY

SPSS Inc. Released (2017) Statistical Package for Social Sciences. IBM SPSS Statistician for Windows, Version 20.0

United Nations. (2015). - *SDG indicators. United Nations Statistics Division*. Retrieved April 19, 2022, from https://unstats.un.org/sdgs/report/2019/goal-04/