The Development of Reading Assessment Manager for Grades 2-6 Pupils of a Progressive School in Metro Manila, Philippines

Sara Jane Sanclaria - Cruz
School of Graduate Studies, Pasig Catholic College, Pasig City, Metro Manila, Philippines

Abstract
In the recent 2018 Programme for International Pupil Assessment (PISA), the Philippines ranked lowest in reading comprehension among 79 countries. It is an alarming result that made the Department of Education further strengthen its program on helping non-readers. True as it may seem that the non-readers should be given focus, on another note, one should also consider that part of this problem are pupils who can read (decode) but has low comprehension. This is a gap that the department seem to overlook. Few to no program is being implemented for pupils who can read but their comprehension is not at par to their grade level. Thus, the researcher developed Reading Assessment Manager for the school which stores an inventory of pupil’s reading progress and recommends level appropriate books to further development of their skills. Using Type 1 Developmental Research Design, the researcher managed and evaluated the whole creation of the system from its analysis, design, development, and evaluation. Participants of the study are the 12 pupil-participants, 5 Language Arts Teachers, 3 Coordinators, and 2 Software Engineers. Administered System Usability Scale (SUS) showed that the developed Reading Assessment Manager (RAM) outperformed 90-95% systems in the industry. Likewise, it garnered an Above Average user satisfaction from the respondents in the given Post Study System Usability Questionnaire (PSSUQ). A paired sample t-test also showed that the Language Arts teachers and Coordinators have rated the Automated Reading Test Administration using RAM higher than the Manual Reading Test Administration with a significance p-value of 0.010.

Keywords: reading assessment software, reading comprehension
1. Background of the Study

In the global industrial scene, this generation has now entered a new age, called the 4th Industrial Revolution or also known as the Industry 4.0. According to McGinnis (2018), this industry is a way of describing the blurring of boundaries between the physical, digital, and biological worlds. This concept of Industry 4.0 started in Davos, Switzerland, when the World Economic Forum happened in 2016. Klaus Schwab, the founder, and executive chairman of Geneva-based World Economic Forum, coined the term the 4th Industrial Revolution (WEF, 2016).

Industry 4.0 is an interconnectedness of technological advances such as artificial intelligence (AI), robotics, the Internet of Things (IoT), genetic engineering, quantum computing, and other technologies. It refers to how technologies are interacting and transforming human’s physical lives by smoothly merging with everyone’s reality. (McGinnis, 2018)

In the advent of COVID-19, wherein social distancing and human interaction is highly discouraged, we have seen what Industry 4.0 can offer to the people. With these, one might notice that technology is redefining the types of jobs that people would have in the future. Machines and technologies would somehow replace the human workforce whether we like it or not, unless an individual possesses the soft skills that these technologies can never replace.

With the rise of Industry 4.0, where technology predominantly automate almost every manual task, curiosity, creativity, empathy, problem-solving, communication, and ability to read, comprehend, and interpret numeric and qualitative data are more important than ever. In this changing paradigm in the workplace of tomorrow, workers who possess the aforementioned soft skills can help organizations adapt and compete in ways that machines cannot—which makes these individuals more sought-after than those who lack the soft skills needed (Foutty, 2018).

Being one of the soft skills for Industry 4.0, reading comprehension skills of a person now becomes a predictor of one’s success in the industry. But how are the school-aged children been doing when it comes to their reading skills?

1.1. Global data on reading

According to the statistical report of the United Nations Statistics Division (UNSD, 2015), it was estimated that 617 million children and adolescents of primary and lower secondary school of age does not meet the minimum proficiency standards in reading and mathematics. This students below the reading proficiency standard are highly concentrated in Sub-Saharan Africa with 88 percent of the population lies in it, followed by Central and Southern Asia with 81 percent, Northern Africa and Western Asia with 57 percent; Latin America and the Caribbean with 36%; Eastern and South-Eastern Asia with 31 percent; Oceania with 22 percent; and Europe and Northern America with 14 percent. Globally, 58 percent of the school-aged children are below reading proficiency.

This reading proficiency crisis not just threatens an individual’s ability to bring him or herself out of poverty, but also, puts the economic future of the country into jeopardy.
1.2. Reading skills of the Filipinos

In the national scene, the Programme for International Student Assessment (PISA 2018), ranked the Philippines lowest in reading comprehension among 79 countries. Student representatives of the Philippines scored an average of 340/1000, which is lower than the global average standard rating of the Organization for Economic Cooperation and Development (OECD). OECD is the prestigious international organization that administers PISA every 3 years which aims not just to rank countries, but to take a good look on which country has the best educational system and curriculum, for other countries to learn from them. Currently, China ranks first followed by Singapore, Macao, Hong Kong, and Estonia.

Currently, the Department of Education is further enriching its program on helping non-readers. True as it may seem that the non-readers should be given focus, on another note, one should also consider that part of this problem are pupils who can read (decode) but has low comprehension skills. The result of the OECD is a clear indicator that this is a gap that the department seem to have overlooked. Only few to no program is being implemented for pupils who can read but their comprehension is not at par with their grade level.

One of the solutions that the Department of Education has thought of is to strengthen the implementation of the Philippine Informal Reading Inventory (Phil-IRI) which aims to assess the learners’ reading levels which can in turn help in the creation of specialized reading programs.

Phil-IRI is a classroom-based assessment tool that is used to get a reading inventory of the pupils’ reading levels in terms of their performance in oral reading, silent reading, and listening comprehension. It is designed to tally a reading inventory of pupils’ progress in both English and Filipino reading comprehension (Ganzon, 2018).

1.3. Concerns in the research locale and the relevance of technology

Locally, this concern, specifically in enriching the pupils’ reading comprehension in Filipino, is also an issue for Antioch School Manila, the research locale of the study. By aligning the school’s reading program with that of DepEd, it was conceived that it would be cost-efficient and effective if Phil-IRI would also be administered rigorously in the school for it could target both reading skills of the pupils in English and Filipino at a minimum cost.

On the other hand, certain setbacks still challenge the administration of the Phil-IRI in the school. The teachers have expressed that it is time-consuming to administer and is not appropriate for an online set-up. The manual administration takes a lot of time and effort on the end of the teachers since this should be done one-on-one with every student. And with the new modality of teaching because of the pandemic; where face-to-face is not possible, administration of the Phil-IRI cannot be conducted and needed to be shelved for the meantime.

But with the changing world order brought by the advances in technology, educational leaders should know how to use technology for their own gain and meaningful purpose. Having an abundance of technology can help school leaders innovate the way schools offer interventions and programs through automation. School leaders should understand that technology serves as a means to an instructional end and how to create the conditions that foster such uses (Dexter, 2018).
Now that technology is available and it can help school administrators design an innovative way to continue learning and program interventions, especially in the time of COVID-19, schools can use the created Phil-IRI and migrate it online for it to be able to reach the pupils even at the comfort of their homes. Hence, it is from this standpoint, that the researcher created Reading Assessment Manager (RAM) for Grades 2-6 Pupils of a Progressive School in Pasig City, Metro Manila.

This Reading Assessment Manager generates an inventory of pupil’s reading progress and recommend appropriate books from the system’s database to further develop pupil’s reading comprehension skills. The web-based Reading Assessment Manager posed as a great help especially in this time of pandemic wherein face-to-face engagements are not allowed. Through undertaking this designing and developing of the Reading Assessment Manager, school administrators build a strong sense of contribution to the automation of the school programs which is relevant in the current situation.

2. Development of the Reading Assessment Manager (RAM)

In this time of pandemic wherein face-to-face interaction is prohibited, reading programs become hard to administer since mostly, it is done manually and face-to-face. But because of the availability of technologies, now it becomes feasible to do a computer-aided reading comprehension assessment. Currently, only few in the likes of Scholastic Inc., Adarna Publishing, and C&E Publishing offer online reading programs for Philippine schools. These two reading programs and online platforms were used as bases for the framework of the Reading Assessment Manager that the researcher created. Following the system flow of SOAR wherein the pupils will (1) choose their reading genre interest; (2) answer reading assessments to bring about individualized reading level; (3) get a book recommendation list based on the chosen genre and reading level result, the Reading Assessment Manager was developed using the same system flow. An additional feature of the system is the reading and comprehension assessment and book list recommendation in English and Filipino.

3. Methodology

This research utilized the Developmental Research Design in creating and evaluating the developed Reading Assessment Manager. The Developmental Research Design according to Richey (1994), is a “systematic study of designing, developing, and evaluating instructional programs, processes, and products that must meet criteria of internal consistency and effectiveness.” It has two types which is predominantly different from each other in terms of the chosen emphasis and focus on the product. Table 1 below shows the difference summary of the Type 1 and Type 2 Developmental Research.
Table 1. Summary of the 2 Types of Developmental Research

<table>
<thead>
<tr>
<th>Emphasis</th>
<th>Product and Tool Research or Type 1 Developmental Research</th>
<th>Model Research or Type 2 Developmental Research</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study specific product or program</td>
<td>Study of design, development, or evaluation processes,</td>
<td>New design, development, and evaluation</td>
</tr>
<tr>
<td>development, &amp;/or evaluation project</td>
<td>processes, tools, or models (can focus only on ONE phase)</td>
<td>procedures &amp;/or models and conditions that</td>
</tr>
<tr>
<td></td>
<td></td>
<td>facilitate their use</td>
</tr>
<tr>
<td></td>
<td>Lessons learned from developing specific products and</td>
<td>CONTEXT-SPECIFIC CONCLUSION</td>
</tr>
<tr>
<td></td>
<td>analyzing the conditions that facilitate their use</td>
<td>GENERALIZED CONCLUSION</td>
</tr>
</tbody>
</table>

Source. Ibrahim (2016)

In this study, the research was guided by the Type 1 Developmental Research Design. Aimed to develop a product in the form of the Reading Assessment Manager, to learn the systematic process of developing it, and to analyze the appropriate condition to maximize the system’s usability, the researcher chose Type 1 Developmental Research Design.

Following this methodology, the research underwent the (1) analysis; (2) design; (3) development; and (4) evaluative phases of the Type 1 Developmental Research Design (Ritchey, 1997) and integrated the stages of Rapid Prototyping, an instructional system design, for a more robust and systematic formative evaluation process.

4. Results
4.1. Usability and learnability evaluation using System Usability Scale (SUS)

The Reading Assessment Manager (RAM) was evaluated by 22 respondents, comprised of 12 pupil-respondents, five (5) Language Arts Teachers, three (3) School Coordinators, and two (2) Software Engineers, using the System Usability Scale (SUS). SUS is a type of User Acceptance Test that aims to evaluate the non-functionality feature of the system (Sauro, 2018). Specifically, it aims to test the ease with which the user can learn, operate, prepare inputs and outputs, through interaction with the system. Hence, a system beta-testing was done before the administration of the SUS.

Analyzing the results of the SUS administered to the 22 respondents, it was revealed that over-all, the system has garnered a SUS Score of 81.6 which can be interpreted as with excellent usability. This excellent usability score of the Reading Assessment Manager can further be rated with a Grade of A which indicates superior performance according to Sauro (2018), and in turn can be considered as acceptable according to Bangor, et.al (2008).

Subsequently, using the curved grading scale interpretation of SUS score developed by Sauro (2011), as given in Table 2 below, reveals that the SUS Score of 81.6 also falls in the percentile range of 90-95.
Table 2. Curved grading scale interpretation of SUS scores

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

Source: Sauro (2011)

According to Sauro (2018), SUS as developed by John Brooke in 1986, have collected 30 years of usage and data from over 10,000 responses and hundreds of products (software apps, hardware, mobile apps, websites, or voice user interfaces). It is from here that the SUS percentile rank is being compared to. Thus, the percentile range of 90-95 of the Reading Assessment Manager explains that it scores better than the 90-95% of the scores in the database.

4.2. User satisfaction results as revealed by the Post-Study System Usability Questionnaire (PSSUQ)

Just like SUS, the Reading Assessment Manager was evaluated by 22 respondents comprised of 12 pupil-respondents, 5 Language Arts Teachers, 3 School Coordinators, and 2 Software Engineers after the beta-testing and walkthrough using the Post-Study System Usability Questionnaire (PSSUQ).

PSSUQ is a 16-item non-functionality usability questionnaire similar to SUS. It was developed at the IBM Design Center in 1992 and designed as a scenario-based usability test for a more targeted evaluation (Rotolo, 2017). Specifically, it assesses the system’s usefulness, information quality, and interface quality. For this research, the researcher administered the PSSUQ on top of the SUS for a more thorough evaluation of the Reading Assessment Manager.

Analyzing the results of the PSSUQ, revealed that the Reading Assessment Manager garnered a mean score of 2.65 for System Usefulness, 2.92 for Information Quality, and 2.38 for Interface Quality. This indicates that the three (3) sub-scales received an Above Average rating from the respondent-evaluators. And among the sub-scales, it is the Interface quality that received the highest rating, for in PSSUQ, the lower result the better. Likewise, the system evaluation using the PSSUQ also gathered an over-all result of 2.67 which can be interpreted as with Above Average rating since it is below the 4.00 PSSUQ score average threshold. According to Lewis (2002), the lower the PSSUQ score, the better the result.

On the other hand, looking at the question items closely would give a more in-depth understanding of the area that should be improved in the system. Question item number 7 which is “The system gave error messages that clearly told me how to fix my problems,” got a PSSUQ score of 4.29 which can be interpreted as Below Average. This means that more effort should be made in designing a good error message.
Going further than the surface of the mean, the researcher also conducted a multiple regression analysis to see which of the three (3) sub-scales has the greatest effect on the overall result of the PSSUQ. Looking at the results closely, it was revealed that Interface Quality with a B coefficient of .250, has the weakest relation with the overall user satisfaction of the evaluators and has the lowest score among the three (3) sub-scales as per the PSSUQ.

Embarking from this, it can be said that among the three (3) sub-scales, System Usefulness, Information Quality, and Interface Quality, it is in the Interface Quality that the researcher needs to perform further developments. This claim was also strengthened by the respondent’s focus group discussion responses and questionnaires wherein they recommended to add more illustrations or graphics that can engage and motivate the students while using the Reading Assessment Manager (RAM). Specifically, 2 of the 10 respondents expressed that it would be better if there would be more pictures, especially for the reading passages. Likewise, the font style a bit small for 50% of the respondents. Thus, further revisions should be made in these areas to improve the Interface Quality of the Reading Assessment Manager.

4.3. 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 resource 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. In here, 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 explain to us that Automated Administration outperformed the Manual Administration of the Reading Assessment.

4.4. Participants’ reflections after being part of the whole research process

The five (5) Language Arts teachers and three (3) school coordinators were interviewed in a focus group discussion to share how they felt after being part of the whole research process, from analyzing, designing, developing, and to evaluating the Reading Assessment Manager.

In the interview, it was shared that the participants felt happy and proud for being part of the research process and for being the first one to use the system. They said that a system such as this is really needed by the pupils of Antioch School Manila.
4.5. Deep insights on Educational Leadership practice that may be derived from the study

a. It is essential that educational leaders know how important Participative management is in schools for it gives the stakeholders the opportunity to contribute and be part of the school improvement efforts and initiatives. Likewise, it gives the stakeholders that sense of accountability and ownership of the program which engages and motivates them more.

b. Teachers and academic coordinators play distinct important roles in ensuring that programs and interventions are addressing what it aims to address. This is the main reason why the Instructional Goal Setting questionnaire was only given to the coordinators because they have the final say on these matters. On the other hand, establishment of the program system requirements were given to both teachers and coordinators because they know what curriculum content should be included and how it should be presented.

c. Data-driven interventions strengthen supports to claims and concretize program directions since it is backed up by data evidence and systematic methods of decision-making.

d. It is perennial especially in our current scenario, that educational leaders employ flexible and innovative solutions to in order to address the learners’ individualized needs that can overcome challenges brought by the pandemic and other pressing educational as well as societal issues and concerns.

5. Conclusion

From the above findings, the following conclusions are made.

1. The Reading Assessment Manager developed performs 90-95% higher than the other systems in the industry.

2. The respondents rated the Reading Assessment Manager with an Above Average user satisfaction.

3. Interface Quality with a B coefficient of .250 has the strongest relation with the overall user satisfaction of the evaluators and has the lowest score among the three (3) sub-scales.

4. Question item number 7 in the PSSUQ got a score of 4.29 which can be interpreted as Below Average.

5. There is a significant difference between the manual test administration and pupils’ reading data storage vis-à-vis the automated in terms of needed user skills and resource allotment as perceived by the teachers.

6. Automated administration of the reading assessment using RAM outperformed the manual administration of the reading assessment.

7. 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.

8. As perceived by the teacher respondents and as stated during the focus group discussion, having such a Reading Assessment Manager that is available online and off-line, shows how responsive and aligned the system is to the current scenario and needs of the pupils.
9. It is essential that educational leaders see the importance of educational technology especially in the current set-up as brought by COVID-19. Now, more than ever, school leaders need to strengthen their technological departments so as to employ flexible and innovative solutions in order to adopt to the changing times.

10. The whole research activities from environmental and situational scanning through Front-end Analysis; to analyzing the prevailing needs of the school through a Needs Assessment Survey; to designing the intervention (the Reading Assessment Manager) based on the analysis and Instructional Goals set by the coordinators; to the development of the system using the given system requirements; and finally, to the evaluation of the system using usability tests and focus group discussions, shows the support these data gives and the systematic way of addressing the needs of the end users and stakeholders in concretely addressing issues, concerns, and challenges that ensure data-driven individualized quality education is being provided to each of the pupils.

11. Data-driven interventions of educational leaders are important since it reflects that decisions and judgements made were backed up by data and evidence which serves as the foundation in ensuring that such interventions are addressing what it needs to address.

12. From an educational leader’s point of view, participation of the school stakeholders in the creation of school programs such as this is a meaningful experience that can further engage each member in the project, as revealed by the focus group discussion conducted with the teacher respondents.

13. Participative management in schools gives stakeholders the opportunity to contribute to school improvement efforts and initiatives.

6. Acknowledgment (TNR 12pt. bold)
This paper is a thesis submitted in partial fulfillment of the requirements for the researcher’s Master of Arts in Education-Major in Educational Management degree. It wouldn’t be completed without the expertise of Mr. Sammuel P. Sanclaria, Dr. Perlita Antonio, Sir Randy Apolinario, Mrs. Marlyn Tolosa, and Dr. Rosaminda Valdez.

The researcher also acknowledges the participation of the Directress, Coordinators, Language Art Teachers, and Pupils of Antioch School Manila. Their participation is the core of this study. In addition, the researcher is grateful for the time and guidance given by the Software Engineers of the Digital Room Philippines and for the approval of Adarna Publishing for the inclusion of their Children’s Book Titles in the Reading Assessment Manager (RAM).

Above ground, the researcher is in debt to her family, especially to her husband, Germellee Michael Cruz, for being her champion and source of unwavering support, and to God Almighty for the blessings of wisdom and knowledge.

7. References


[34] SOAR (Scholastic online reading Assessment. (n.d.). Retrieved from https://education.scholastic.ca/category/SRI_READING_INVENTORY