

TESS: A Computerized Oral Reading Assessment System for incoming Grade 7
students

A Special Problem Proposal
Presented to
the Faculty of the Division of Physical Sciences and Mathematics
College of Arts and Sciences
University of the Philippines Visayas
Miag-ao, Iloilo

In Partial Fulfillment
of the Requirements for the Degree of
Bachelor of Science in Computer Science by

DIESTRO, Kimberly
SALIBIO, Alexa Therese

Francis DIMZON
Adviser

January 28, 2025

Contents

1	Introduction	1
1.1	Overview	1
1.2	Problem Statement	2
1.2.1	General Objective	2
1.2.2	Specific Objective	2
1.3	Scope and Limitations of the Research	3
2	Review of Related Literature	4
3	Research Methodology	7
3.1	The Slosson Oral Reading Test	7
3.1.1	Reading Level	8
3.2	Description and Functionality of TESS	8
3.2.1	Voice Functionality and Collected Voice Data	9
3.3	Developing TESS	10
3.4	GUI Mockup	12
3.5	Testing the Program	14
3.5.1	Software Evaluation Form	14

3.5.2	Focus Group Discussion Outline	16
3.5.3	Consent Forms	16

Chapter 1

Introduction

1.1 Overview

Iloilo City National High School (ICNHS) is one of the few schools that has a dedicated remedial reading program. It is a program that aims to address the learning gaps of students with reading difficulties. Students with such difficulties are frequently referred to as ‘frustrated’ or ‘struggling’ readers and is defined by a student’s reading level falling short of their grade level. Estuya defines a frustrated reader as someone who “struggles with reading despite appropriate instruction” (Estuya). The reading level of a student is determined through testing, which the program does at the beginning of the year in order to identify students in need of remediation. Once identified, they are grouped together in a section or two so that remedial reading classes can be easily scheduled in hopes of improving their reading levels. Activities are also arranged to encourage students to persevere despite their circumstances.

Incoming students are assessed using the Slosson Oral Reading Test (SORT) to determine the grade level of their word recognition skills. SORT evaluates a student’s word recognition level and grades them into 10 grade levels; wherein the first level P means primary, and the last grade is high school. Grades one to eight serve as the in-between reading levels. Students are asked to read from a set of 10 word lists that correspond to each of the reading levels starting from the lowest reading level and moving upward. A qualified teacher administers the test on a one-on-one and face-to-face basis. (Slosson Educational Publications, Inc.)

The test concludes when the student finishes all levels or incorrectly reads 10 words consecutively. Their reading level is then determined by the total number

of words read correctly divided by 20. Those students whose reading level is determined to be below level 2 are considered in need of intervention and are included in the program for remediation.

Every year, there are over 1200 Grade 7 students enrolled in ICNHS. However, they have narrowed down the testing population to incoming students who have an average grade of below 80% (over the pandemic they have increased it to 83% and below). Still, every year there are at least 300 students requiring assessment. In conducting this test, they identify over 60 students with reading difficulties every year. With this volume of students to be assessed, the process requires multiple teachers to volunteer their time. This is all on top of their other enrollment responsibilities. Due to this, only about 6 teachers volunteer every year in a span of a couple of days.

Students with reading difficulties face more challenges in their education than their peers. Thus, it is important for the school to recognize these issues early on and work with the students and parents to address these learning gaps.

1.2 Problem Statement

In order to better meet the testing requirements for the incoming Grade 7 population, the Committee of ICNHS' Reading Remediation Program is looking into technology that can help them streamline their testing process and improve the rate at which they can assess the population. The development of a Computerized SORT that meets their specifications would greatly help and improve their assessment process.

1.2.1 General Objective

To create a program that could conduct the Slosson Oral Reading Test that meets the testing requirements and specifications of ICNHS Remedial Reading Program. The program should also allow for more students to take the test simultaneously as a group.

1.2.2 Specific Objective

The specific objectives of this research is as follows:

1. To convert SORT into a working computerized assessment.
2. To develop an easy-to-use program that can be used by teachers as-is.
3. To develop a program that can be used offline and in a school computer lab.
4. To identify voice recognition software that can reliably transcribe words spoken by a non-english speaker in a school environment.

1.3 Scope and Limitations of the Research

TESS is predominantly a computerized testing tool and will only generate the result of the Computerized SORT for use by the program's faculty.

TESS will only be launched and tested within Iloilo City National High School (ICNHS). The test group will consist of a select group of students with differing reading levels from the school's Grade 7 population.

Chapter 2

Review of Related Literature

Reading is one of the basic and most essential literacy skills required to be successful in today's world of technology and information. Reading proficiency can affect not only a student's ability to perform in school, but also their self esteem and their attitude towards schooling. This in turn can severely impact the student's future prospects and social mobility. Therefore it is important to ensure that a student is able to achieve the reading milestones / achievements set for their grade.

DepEd's Program "Every Child a Reader", which was started S.Y. 2002-2003, mandates that every pupil shall not proceed to the next grade level unless they have achieved mastery over the basic literacy skills of their grade level; however, despite this and the many efforts of DepEd to encourage and develop the reading skills of the nation's learners, there are still slow and even non-readers in incoming 7th grade students. (Sancada (2022)). Additionally in the 2018 Programme for International Student Assessment (PISA), Philippines ranked last in the world in terms of reading and improved slightly in the 2022 PISA now ranking 6th lowest in the world instead. (PISA, 2018; PISA, 2022; EDCOMM II year one report, 2024)

There is much research and evidence that supports that early intervention, such as during elementary school, is effective at helping slow and frustrated readers improve, catch up, and even excel beyond the literacy milestones of their grade. (Lee et al. (2012)) Failure to intervene early can lead to the student developing poor reading skill which in turn leads to poor academic performance, creating consecutive experiences of failure that can further demotivate students to learn. (Fosudo, 2010 as cited in Sancada, 2022)

However, even if early intervention was not possible, remediation during their adolescence or in later grades, such as high school, is possible and effective (as an example, Rafanan & Raymundo, 2024; Lovett et al, 2021). As part of DepEd's efforts to improve the literacy of pupils and students in the Philippines, many schools have established Remedial Reading Programs in an effort to identify students in need of help. A study aimed to test the efficacy of remedial reading programs in helping struggling readers of Grade 7 students of Bolo Norte High School, found that intervention had a substantial impact on student's reading comprehension. Students were able to improve after a 5-month period of remedial classes, highlighting the positive impact of educational intervention. (Abergos et al. (2024))

Despite the efficacy and the potential of these reading programs, there can still be a lack of support for them. A study looking into the challenges and effective practices of Remedial Reading Programs in Iloilo found that one problem that these programs face is a lack of support from the Parents, School, and even the Local Government Units. A lack of support could mean not enough attention from the parents to support their children, a lack of funds to purchase books and equipment, a lack of interest from teachers to participate and proctor in the program and even a lack of interest from the school to support the program due to it not being a flagship program of the school, which can result in a lack of classroom space and no available schedules to conduct the program. It follows that schools with successful and more effective programs are ones that have the support needed to have proper implementation. (Sancada (2022))

The process of assessing students is the first step to providing support to students in need of remedial reading classes. If these schools cannot efficiently and effectively identify students in need, then the reading programs will not be able to perform to their full potential. Currently, reading assessments are conducted one-on-one with teachers' having to proctor and assess the student's results manually and all on their own.

For example, every year, as mentioned above there are over 1200 Grade 7 students enrolled in Iloilo City National High School. However, the school doesn't have the manpower to administer the test to every incoming student. Even after narrowing down potential students in need, they identify over 60 students with reading difficulties every year. With this volume of students to be assessed, the process requires multiple teachers to volunteer their time. This is all on top of their other enrollment responsibilities. Due to this, only about 6 teachers volunteer every year in a span of a couple of days.

This lack of available staff and resources proves to be a significant bottleneck in being able to identify students in need. One such solution created to streamline the

process of student assessment is to make use of current and emerging technology to create computerized assessments. In a 2023 study by Auphan et. al., they tested the efficacy of using computer-based assessments (CBA) for reading and found that such tools provide a low-cost and easy-to-use tool for professionals to assess reading skills. They even further suggested that CBA's have many advantages and developing tests to make use of the technological advantages of CBA can further improve the efficacy of a reading assessment.(Auphan et al. (2020))

In 2022, All Children Reading (ACR) Philippines, in coordination with the United States Agency for International Development (USAID) and DepEd, developed a Computer-Based Reading Assessment (CoBRA) Pilot and tested the program in 42 schools in the Philippines. CoBRA would allow students to independently take the Philippine Informal Reading Inventory (Phil-IRI) – a standardized reading assessment used as the primary means to determine the reading levels of students based on set benchmark scores. CoBRA would also provide them with instant feedback and assessment through an automated grading system and also used voice recognition software to facilitate the test – mainly Poodll and also Google Speech Recognition for the Filipino language.

The pilot itself was conducted on school premises and made use of the school's available equipment. Ultimately, the pilot was successful with positive reception from students and teachers alike. A teacher was quoted to be saying “it saved them weeks of calculating and counting words, and at same time saved the school from printing papers.” CoBRA offered a more efficient and streamlined way to assess students – although its automated grading system and the voice recognition software used was not accurate and still required significant human intervention to be able to produce accurate results.

A planned second pilot for CoBRA would not proceed as planned for a few reasons. (1) The sponsoring unit with DepED ICTS was dissolved during a transition in leadership. (2) Licensing costs for the software escalated to what were deemed unsustainable levels.

One can learn a few things from the CoBRA pilot. (1) Secure an accurate voice recognition software that can operate in a school setting as well as detect accented english. (2) The program itself must be able to stand on its own or at least be developed using software that schools or the reading program can reasonably sustain.

In regards to voice recognition software, there are a few that come to mind. VOSK, Whisper, and Wav2Vec. All are open source voice recognition software that can be used to develop a computerized reading assessment system.

Chapter 3

Research Methodology

3.1 The Slosson Oral Reading Test

The Slosson Oral Reading Test (SORT) is a quick screening test used to determine a student's reading level. The SORT is composed of a reading list of 200 words grouped into 10 groups of 20. Each group is ordered according to 10 reading levels (P or primary, Grade 1 to 8, and High School). The word list was compiled from standardized school readers, and the reading level represents the median school achievement.

The test is administered one-on-one between the student and the teacher. The student is provided a copy of the word list and is instructed to read down each list loudly and clearly. A student is allowed to skip a word if they do not know how to read it. As the student reads, the teacher is tallying down the number of words read correctly and incorrectly. The goal of the test is to find the starting list, stopping list, as well as the raw score of the student.

The starting list is the reading level list that the student can read all words correctly, while the stopping list is the list the student will get all 20 of the words incorrectly. Lastly, the raw score is the total number of correctly-read words throughout the whole list. This raw score is divided by 20 to then get the student's reading level.

3.1.1 Reading Level

In addition to the reading level included in SORT, the reading level used in the Philippine Informal Reading Inventory (PHIL-IRI) will also be used to identify reading level groups for the testing of TESS.

The PHIL-IRI 2018 manual categorizes reading level as the following:

1. Independent - Readers are independent and can read with almost perfect oral reading and comprehension
2. Instructional - Readers in this level benefit the most from instruction and direction from the teacher.
3. Frustrated - Readers in this level find reading material so difficult that they cannot interact with the material.

3.2 Description and Functionality of TESS

TESS is a computerized reading assessment program that will allow students to take the SORT with minimal to no assistance from a teacher. The following steps described were specifically requested and specified by the faculty-in-charge of ICNHS' reading remediation program.

In this computerized SORT, students will read out loud from a randomized list of words flashed to them, word-by-word on the screen. This list of words is predetermined and is categorized by reading level. If the student is deemed correct, the program will proceed with the next word in the list. If the student was incorrect, then the program will remain at the current word. The student is allowed to attempt the word three times, with a timer of 10 seconds per attempt.

Additionally, the faculty wants to decrease the amount of stress and frustration the examinee will experience during the assessment, so TESS includes a built-in timer for each item in the test. Once eleven (11) seconds have passed an item will be skipped, and once ten words in a row have been skipped, the test will end and the examinee's reading level will be calculated.

A progress bar will also be visible in the bottom, so that students are able to track their progress through the assessment.

The program will also evaluate the students and provide the teacher with an

excel sheet for each student containing their results, which includes an itemized list of which words in the SORT they got right and wrong.

The test portion of the program is voice activated, and the examinee can make use of certain keywords to navigate through the test. Such as START to start the test; SKIP to skip a word; and STOP to end the test prematurely.

TESS should be able to run on an offline computer with a headset.

3.2.1 Voice Functionality and Collected Voice Data

TESS' voice functionality is the most important part of the program. The speech-to-text program included should be able to function in a school environment and also be able to accurately assess accented English.

Student's voice will be recorded by the program, assessed, and then discarded. The program does not need to store data of any examinee's voice to function.

Other Collected Data

TESS will take note of the following information:

- Student Name
- Student's School
- Proctor's Name

This information is used to generate the excel report after the test is finished. TESS does not save this data after the completion of the test.

3.3 Developing TESS

TESS is developed using Python and a voice recognition software. The three main choices for our voice recognition software are the following:

- Vosk,
- Wav2Vec, and
- Whisper

These software were chosen because all three are open source software and do not require any payment to keep it running. The program is also designed to work on a computer with a dedicated headset and microphone.

In order to determine which of the three to use, we developed a TESS prototype that can evaluate the voice input using all three software. We then conducted an informal test to see which of the three is most effective.

However, Whisper was out of the running early on as it was unable to be run quickly enough to use in real time. This left Vosk and Wav2Vec to be tested.

After further research we decided to use Vosk. Vosk was chosen because of the following advantages:

1. Offline Capability,
2. Lightweight and Fast,
3. Customizable Models,
4. Open-Source, and
5. Cost-Effective.

Although, Vosk does have a few disadvantages. Most notably, it has difficulty dealing with the following:

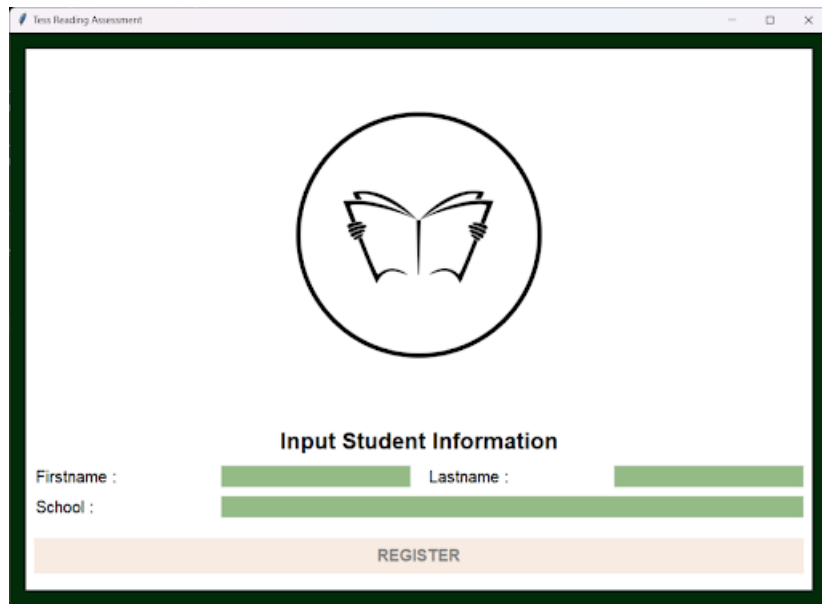
1. Accuracy with Complex Speech,
2. Limited Support for Certain Accents, or Dialects
3. Limited Context Understanding.

In order to further improve Vosk's accuracy in dealing with complex speech as well as to deal with accented English, Jellyfish was used in the post-processing. Jellyfish compares two strings to see how phonetically close they are.

The program is also designed to work on a computer with a dedicated headset and microphone.

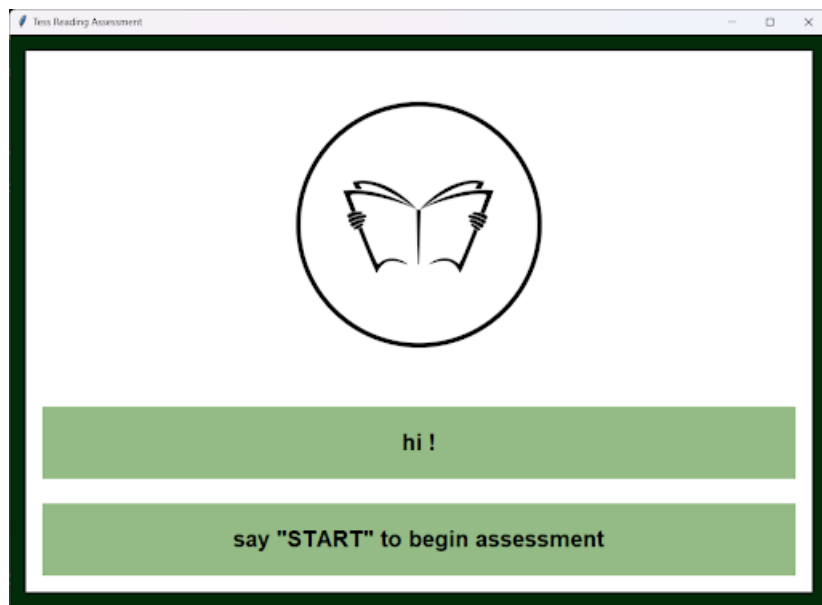
3.4 GUI Mockup

The following are images of the current GUI of Tess.



The mockup shows a window titled "Tess Reading Assessment". At the top center is a circular logo containing a stylized open book with hands holding it. Below the logo, the text "Input Student Information" is centered. Underneath, there are three input fields: "Firstname :" followed by a green box, "Lastname :" followed by a green box, and "School :" followed by a wider green box. At the bottom, there is a wide, light orange button labeled "REGISTER".

Figure 3.1: Student Input Screen.



The mockup shows a window titled "Tess Reading Assessment". At the top center is the same circular logo as in Figure 3.1. Below the logo, there are two green rectangular buttons. The top button contains the text "hi !". The bottom button contains the text "say 'START' to begin assessment".

Figure 3.2: Pre-test Screen.

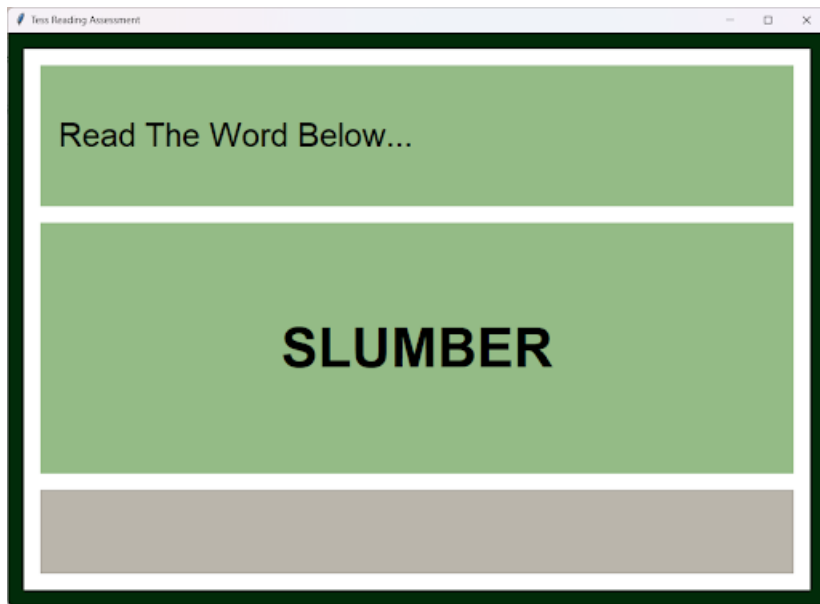


Figure 3.3: Test Screen.

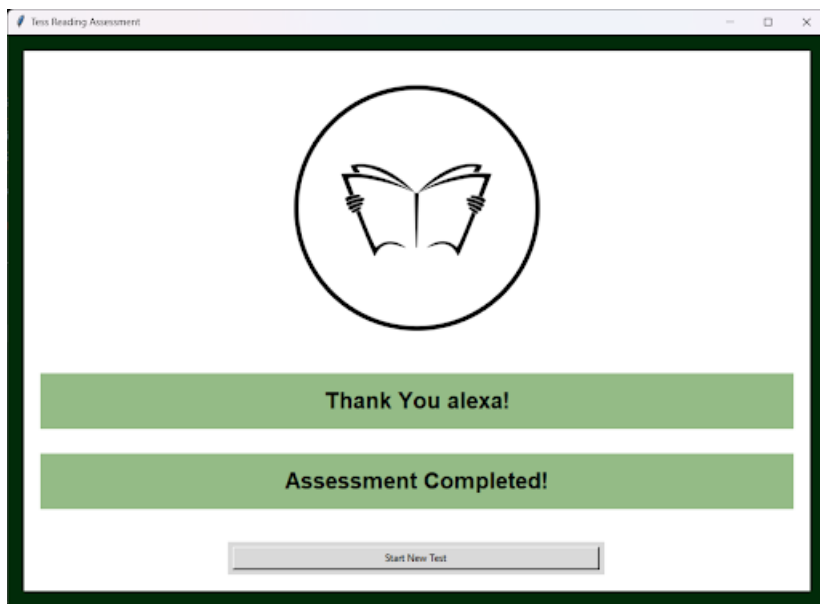


Figure 3.4: Post-test Screen.

3.5 Testing the Program

TESS will be tested at the Iloilo City National High School with a group of six Grade 7 students. Two students of each reading profile (independent, instructional, frustration) will be chosen for the final group of six.

The group will take the manual SORT and the computerized SORT through TESS at the same time – this is to avoid practice bias. The testing duration will be timed, and the results of the tests will be used to determine the efficacy of TESS in relation to a manual SORT.

After the testing, a focus group discussion will be held to determine what the user experience was like.

TESS will then be compared with Manual SORT along the following metrics:

- Speed - how fast does each student complete their assessment?
- Accuracy - how accurate are the results of each student?
- Ease of Use - how easy was the test to be administered / taken?

The Speed and Accuracy metrics can be determined using the information gathered during the assessment. However, the Ease of Use metric will be derived from a Focus Group Discussion composed of participating students and faculty.

3.5.1 Software Evaluation Form

To evaluate how effective TESS is along the Speed, Accuracy, and Ease of Use metrics, we have created an evaluation form.

While the main categories are Speed, Accuracy, and Ease of Use, they will be further expanded into relevant questions which will allow the researchers to better gauge TESS' performance in each of those metrics.

For each metric and its corresponding descriptive characteristic, TESS will be graded on a scale of 1 to 5; with 1 being “Poor” and 5 being “Very Good”. The scale is as follows, from least to greatest.

1. Poor (P)
2. Fair (F)
3. Average (A)
4. Good (G)
5. Very Good (VG)

Questions	P (1)	F (2)	A (3)	G (4)	VG (5)	Comments (Notes)
SPEED						
(1) Program loads each word quickly and without delay.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
(2) Program processes input without significant delay.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
(3) Program finishes the SORT assessment quickly.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
ACCURACY						
(1) Program accurately transcribes the spoken word.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
(2) Program can identify successful and unsuccessful word readings.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
EASE OF USE						
(1) The interface is easy to understand and use.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
(2) The instructions provided are easy to understand and follow.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
(3) The functionality of the program is easy to explain and teach..	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

Figure 3.5: Software Evaluation Form.

3.5.2 Focus Group Discussion Outline

The focus group discussion will center around the experience of the faculty and students while using TESS. The discussion will last for around 30 minutes. The following is the set of questions we will use as a guideline for the discussion.

1. What was the experience of using TESS to take the reading test like?
2. How did using TESS differ from taking / administering SORT manually?
3. What difficulties did you encounter while using TESS?
4. Among the TESS models, which one did you like using most?
5. If you could add one feature or function to TESS what would it be?

3.5.3 Consent Forms

Students participating in the TESS testing will be required to fill out parental consent forms in order to participate. There are two forms. One in English, and one in Hiligaynon.

References

- Abergos, L. I., Cruz, J. R., Lasala, J. C., Prado, S. N., Tapar, P. K., Cañeza, D. C., & Ocampo, D. M. (2024). Effectiveness of remedial reading to struggling readers of grade 7 students. *SHS Web of Conferences*, 182, 01004. doi: 10.1051/shsconf/202418201004
- Auphan, P., Ecalle, J., & Magnan, A. (2020, Sep). High potential of computer-based reading assessment. *Canadian Journal of Learning and Technology*, 46(1). doi: 10.21432/cjlt27847
- Department of Education, Pablo, F. E., Diaz, L. E., Ongtengco, M. H. P., Alcazar, M. Y. C., Salvador, A. M. M. S., & Nava, F. J. G. (2018). *The Philippine Informal Reading Inventory Manual 2018*. Department of Education.
- Estuya, J. A. (2024). *Home and school based-factors affecting the reading performance of frustration-level readers* (Unpublished doctoral dissertation). Columban College, Inc.
- Lee, G.-L., Gable, R., & Klassen, V. K. (2012). Effective reading remediation instructional strategies for struggling early readers. *Procedia - Social and Behavioral Sciences*, 46, 822–827. doi: 10.1016/j.sbspro.2012.05.206
- Lovett, M. W., Frijters, J. C., Steinbach, K. A., Sevcik, R. A., & Morris, R. D. (2021, May). Effective intervention for adolescents with reading disabilities: Combining reading and motivational remediation to improve outcomes. *Journal of Educational Psychology*, 113(4), 656–689. doi: 10.1037/edu0000639
- Rafanan, J. A., & Raymundo, J. L. (2024, Jul). Impact of customized reading remediation and intervention in reading fluency of high school students. *International Journal of Linguistics, Literature and Translation*, 7(7), 66–74. doi: 10.32996/ijllt.2024.7.7.8
- Sancada, R. T. (2022). Remedial reading program in the schools division of iloilo: Challenges and effective practices. *Psychology and Education: A Multidisciplinary Journal*, 1–19. doi: 10.5281/zenodo.6957238

Second Congressional Commission on Education. (2024). *Miseducation: The failed system of philippine education, EDCOMM II year one report* (Tech. Rep.). Second Congressional Commission on Education.

Slosson Educational Publications, Inc. (n.d.). *Slosson Oral Reading Test - Revised*. Retrieved from <https://www.slossonplus.com/sort-r3>

United States Agency for International Development. (2022, Nov). *Computer-based reading assessment pilot report* (Tech. Rep.). United States Agency for International Development. Retrieved from <https://shared.rti.org/content/computer-based-reading-assessment-pilot-report>