Instrument Development and Pilot Report

Improving Deaf Children's Reading Through Technology in Morocco

Institute for Disabilities Research and Training
in collaboration with École Nationale Supérieure des Mines de Rabat

Prepared by:
School-to-School International
for All Children Reading: A Grand Challenge for Development

July 2017

Table of Contents

I. Introduction

All Children Reading: A Grand Challenge for Development (ACR GCD)—a partnership between the United States Agency for International Development (USAID), World Vision, and the Australian Government—is an ongoing series of grant and prize competitions that leverage science and technology to source, test, and disseminate scalable solutions to improve literacy skills of early grade learners in developing countries. Round 2 of ACR GCD, which started in 2014 and continues through 2017, supports technology-based innovations to improve early grade reading outcomes in developing countries.[1] These technology-based innovations feature three focus areas:

1. Mother tongue instruction and reading materials
2. Family and community engagement
3. Children with disabilities

The Institute of Disabilities Research and Training, Inc. (IDRT) was selected as a Round 2 grantee in the children with disabilities focus area. Children who are deaf and hard of hearing in Morocco face great challenges learning to read in Modern Standard Arabic (MSA). Schools that serve students who are deaf or hard of hearing in Morocco lack quality teaching and learning materials to support students' acquisition of reading skills. In addition, many students do not have sufficient access or exposure to whole language at home or at school. They are not given sufficient opportunities to develop their Moroccan Sign Language (MSL) abilities, which further exacerbates their reading development. IDRT's Improving Deaf Children's Reading Through Technology in Morocco project is developing software, *MSL Clip and Create*, that will enable teachers to create learning materials that reinforce vocabulary and concepts to support their students' language and reading development. To date, the project has developed 2,200 MSL graphics and video clips to represent 5,500 MSA words with supporting concept graphics and continues to grow.

To understand the impact of technology-based innovations on improving the literacy skills of early grade learners, an assessment of early grade reading skills is conducted as part of each ACR GCD grant project. To measure this, ACR GCD uses the Early Grade Reading Assessment (EGRA) to systematically assess reading skills across all Round 2 grantees including projects working with students with disabilities. Typically, the EGRA is administered as an oral assessment to measure students' most basic foundational skills needed for literacy acquisition in the early grades—specifically, recognizing letters of the alphabet, reading simple words, understanding sentences and paragraphs, and listening with comprehension. The EGRA methodology was developed under EdData II and has been applied in more than 30 countries and 60 languages.

However, to STS's knowledge, very few countries have conducted reading assessments with students who are deaf or hard of hearing. As part of the Moroccan Sign Language Assistive Technology for Child Literacy project and with the help of Moroccan stakeholders, School-to-School International (STS) and IDRT adapted an existing MSA EGRA that was used in Morocco in 2015 to appropriately assess the

---

[1] Retrieved from: http://allchildrenreading.org/about-us/

reading skills of the project's target population. The new tool is called an early grade reading and sign language assessment (EGRSLA).

The EGRSLA and additional language assessment instruments, detailed in section III, were developed over a six-month process that culminated in an adaptation workshop to review and revise the instruments in October 2016. The instruments were piloted immediately after the workshop. A second adaptation workshop to revise and re-pilot the instruments will take place in 2017, and operational baseline data collection will commence with the new school year in 2017. Endline data will be collected at the end of the school year in 2018.

This pilot report provides information about the process of developing and piloting the EGRSLA and additional language assessments for deaf or hard of hearing students in Morocco, as well as results from data analysis. The report also highlights challenges and potential revisions recommended for the next phase of the adaptation process.

II.   Pilot Study Purpose

The purpose of the pilot study was to ensure that the EGRSLA and other language assessments appropriately measure MSA reading and MSL abilities, as well as the contextual factors impacting Grade 1 and 2 students who are deaf or hard of hearing in Morocco.

Originally, the project anticipated conducting one adaptation workshop and pilot test to finalize all instruments, followed by an operational baseline data collection. Several challenges were encountered during the initial adaptation and pilot process:

1. Each of the eight intervention schools use different variations of MSL.
2. Data collection is complex and includes capturing data through tablets, video, and paper.
3. A non-intervention group with similar characteristics to the intervention group was not available on which to rigorously pilot the instruments.
4. There was inadequate stakeholder representation from the Moroccan Deaf Community during the adaptation process.
5. Only two potential assessors with the requisite skills were identified for data collection.

Because of these challenges and delays in the IDRT project rollout, it was determined that the initial baseline data collection would instead serve as a pilot study for the instruments and protocols.

The process for developing and piloting the EGRSLA, language and teacher assessments, and data collection procedures included the following steps:

1. Recruit a reading specialist with expertise working with students who are deaf or hard of hearing.
2. Create a working group consisting of project and Ministry of Education staff and representation from the Deaf Community.
3. Determine which EGRA subtasks are appropriate for students who are deaf or hard of hearing.

4. Explore additional language assessments to determine students' and teachers' sign language proficiency.
5. Review existing EGRA tools from the Moroccan context and other MSA EGRAs.
6. Develop preliminary drafts of instruments using existing tools.
7. Develop additional language assessments to be used for students and teachers.
8. Adapt the EGRSLA and additional language assessments with the working group members.
9. Train two assessors to collect pretest data.
10. Pretest all instruments in two schools serving students who are deaf or hard of hearing but not assigned to the intervention group.
11. Revise the instruments based on feedback and pretest data.
12. Conduct a five-day pilot data collection training for two assessors and one supervisor.
13. Collect pilot data from students who are deaf or hard of hearing in Grades 1 & 2 at eight intervention schools.
14. Review and clean the pilot data.
15. Conduct pilot data analysis.
16. Present preliminary findings to stakeholders in Rabat, Morocco.
17. Produce a pilot report with recommendations for further revisions.

III.   Pilot Instrument Development

To effectively measure changes to the reading skills of Moroccan students in Grade 1 and 2 who are deaf or hard of hearing, an EGRSLA and additional language assessments needed to be developed. STS, IDRT, and ACR GCD collaborated to review existing tools and ensure alignment with the students' learning environment. Prior to the adaptation workshop, STS convened a small working group to determine the core EGRA subtasks that should be included and to explore additional language assessments that could be appropriate for the target population. All instruments were developed to assess MSA reading ability as expressed through MSL, which varies in the different regions in Morocco. In total, three student-level instruments and five school- or teacher-level instruments were developed.

The student-level instruments include
1. EGRSLA
2. Student expressive vocabulary test (EVT)
3. Student questionnaire

The school- or teacher-level instruments include
1. Teacher expressive vocabulary test (EVT)
2. Teacher communication development inventory (CDI)
3. Student communication development inventory (CDI)
4. Teacher questionnaire
5. Classroom observation instrument

Adaptation Workshop

An adaptation workshop was held from October 3 to 7, 2016, to review and revise the EGRSLA and additional language assessments prior to pretesting. The workshop, led by STS facilitators, brought together experts from the IDRT project as well as the Ministry of National Education and Vocational Training (MNVET) in Morocco including:

- Ismael Oughanim, MSL expert and MSL-MSA interpreter (hard of hearing)
- Mohammed Kettani, MSL expert and MSL-MSA interpreter
- Mohamed Anouar Boukili, division head for Children with Disabilities at MNVET
- Hajar Ajmi, teacher of students who are deaf or hard of hearing in intervention school
- Fatna Eddibaoui, teacher of students who are deaf or hard of hearing in intervention school
- Soukaina Mahtouch, teacher of students who are deaf or hard of hearing in intervention school (deaf)
- Mohamed El Hamri, reading expert
- Abderrahim Haracha Ras, curriculum and assessment expert
- Mohamed Ouknider, special education expert and IDRT project coordinator
- Abdelhadi Soudi, IDRT co-principal investigator
- Joshua Josa, disability inclusive education specialist at USAID (deaf)
- Corinne Vinopol, deaf education and technology specialist and IDRT project director
- Amy Lederberg, professor of educational psychology, special education, and communication disorders at Georgia State University (hard of hearing)

EGRSLA and Additional Language Assessments

The EGRSLA is based on an existing MSA EGRA that STS developed for use in Morocco in 2015. Because the MSA EGRA was intended for children with hearing, each subtask's content and administration protocol was reviewed and adapted for MSL during the adaptation workshop. Based on the recommendations of the working group, STS presented seven subtasks to workshop participants for adaptation. Each subtask required the students to demonstrate their MSA reading ability through production of MSL, meaning all subtasks required bilingualism from the participants.

1. **Letter name identification (timed):** This subtask "tests children's ability to recognize the graphemic features of each letter and accurately map it to its corresponding name."[2] In the context of education for students that are deaf or hard of hearing, this subtask is important because fingerspelling is a building block for reading and positively correlates with stronger reading skills.[3] A letter frequency analysis was conducted on early grade reading text. Based on its findings, Arabic letters were listed randomly in a table with some letters presented multiple times according to the frequency with which the letter appears in early grade reading text.

---

[2] RTI International. (2016). Early Grade Reading Assessment (EGRA) Toolkit, Second Edition. Washington, DC: United States Agency for International Development.

[3] Baker, S. (2010, July). Visual Language and Visual Learning Research Brief: The Importance of Fingerspelling for Reading (Learning from Research No. 1). Retrieved from http://vl2.gallaudet.edu/files/7813/9216/6278/research-brief-1-the-importance-of-fingerspelling-for-reading.pdf

Students were asked to provide the name of each letter in the MSL-fingerspelling alphabet.

2. **Diacritic name identification (untimed):** This subtask tested children's ability to recognize eight Arabic diacritic symbols presented in isolation and accurately map them to their corresponding names. In Moroccan mainstream schools, children are taught the names of diacritics when they learn the alphabet; diacritics are included in early grade reading instruction and text since diacritical markings change the meaning of words. There is less consistency in how students who are deaf or hard of hearing learn diacritics. Participants in the adaptation workshop brought to the group's attention that there are regional and school variances in the use of diacritics in fingerspelling. Some regions and schools fingerspell words using diacritics and some do not. There are also different systems across regions for fingerspelling diacritics. As such, this subtask was removed after the pretest.

3. **Syllable identification (timed):** The syllable identification subtask, like the letter name identification subtask, tests children's ability to recognize and name letters. At the time of the 2015 MSA EGRA development, a syllable frequency analysis was conducted on early grade reading text. Based on those results, Arabic letters were listed randomly in a table with some letter and diacritic combinations presented multiple times corresponding to frequency. In its adaption for MSL, this subtask tested children's ability to recognize and fingerspell letter-diacritic combinations.

4. **Familiar word reading (timed):** The familiar word reading subtask presents a list of words that children are expected to be able to read at their grade level and that they have likely encountered before. The subtask consists of short, familiar sight words presented randomly in a table. Word frequency analysis was conducted on early grade reading text to identify a list of familiar words. The adaptation team reviewed the results to ensure the final list included words that (1) were encountered by early grade readers in Morocco, (2) represented different parts of speech, and (3) had corresponding MSL signs. In the context of education for students who are deaf or hard of hearing, this subtask is important as it tests students' ability to read and comprehend written words. The students were required to read the word in written Arabic, understand the word, then translate the word into MSL. Students were asked to produce the corresponding MSL sign for each familiar word.

5. **Unsigned reading and reading comprehension (untimed):** Intended to be a measure of students' ability to read connected text and respond to comprehension questions, this subtask consisted of a short passage and five comprehension questions. Students were given three minutes to read the text to themselves. At the end of three minutes, the text was removed, and students were asked five comprehension questions in MSL regarding the text. To complete this task, students needed to read the passage in written Arabic, understand the word in context, then translate that word into MSL. This subtask was removed after the pretest.

6. **Reading passage (timed) and reading comprehension (untimed) (2 subtasks):** These subtasks measure students' overall reading ability, which is their ability to read connected text, process

connections, relate text to meaning, and make inferences. Referred to as the oral reading fluency subtask on spoken-language EGRAs, students are presented with a short, simple text (around 60 words) and are asked to read the text out loud. For the purposes of this EGRSLA, students were presented with a text in MSA and asked to interpret the text to the enumerator in MSL. As with the EGRA oral reading fluency subtask, students were then asked up to five comprehension questions based on the text they just read. Students were only asked questions for which they had read far enough in the passage to find the answer. The students were required to read the passage in written Arabic, understand the meaning, then translate the meaning into MSL responses on the comprehension questions.

7. **MSL comprehension (untimed):** This subtask is called listening comprehension in EGRAs for hearing students and typically consists of an assessor reading a short passage aloud to the student followed by comprehension questions for the student to answer aloud. For this EGRSLA, students watched a video of a short passage signed in MSL and subsequently responded to videos of comprehension questions in MSL related to the passage. The assessment team determined that the best way to meet reliability standards was to have the passage video recorded, thus ensuring students all experienced the story with the same sign vocabulary. While recording the passage and comprehension questions, great care was taken to use grammar and register appropriate for storytelling for children in MSL.

For all untimed EGRSLA subtasks, students' responses were coded one of four ways: (1) the student provided a correct sign that is in the IDRT software; (2) the student provided a correct sign but not the sign in the IDRT software (a valid regional variation); (3) the student provided an incorrect answer; or (4) the student did not provide a response. Based on guidance from literacy- and deaf-education experts, students were allowed two minutes for all timed EGRSLA subtasks.

In addition to the EGRSLA, the assessment team determined it was important to also capture the students' and teachers' sign language ability with simple vocabulary words and to assess the contextual factors impacting both students and teachers.

- **Student communication development inventory (CDI):** The student CDI is an assessment of expressive vocabulary based on reporting from a teacher or parent. It is frequently used to measure early vocabulary development and has been adapted to many languages, including American Sign Language. In this assessment, teachers and parents are given a list of words and asked to report whether children say or sign them in everyday conversation. For the purposes of this adaptation for students who are deaf or hard of hearing, it was clear that the tool would need to measure students' vocabulary knowledge in MSL of corresponding words in written Arabic. To meet this goal, words for the student CDI were sourced from the IDRT software. The adaptation team then reviewed the list of words for MSL appropriateness for readers in Grades 1-5. The assessment team crossed out inappropriate words (such as words not relevant to the local context or words that would be unknown to early readers) and, if possible, replaced them with appropriate alternatives. The criteria for determining the appropriateness of the word was left to the discretion of the adaptation team.

- **Teacher communication development inventory (CDI):** The adaptation team also decided that teachers should take a CDI to self-report their own MSL vocabulary knowledge. Many teachers of students who are deaf or hard of hearing in Morocco are not fluent users of MSL, and thus, their level of fluency impacts children's access to vocabulary. This is especially true if the child does not have exposure to MSL at home or elsewhere in their daily life. The vocabulary on the teacher and student versions of the CDI were identical.

- **Student expressive vocabulary test (EVT):** This test was originally intended to be an Arabic picture vocabulary test that was designed to measure students' receptive vocabulary skills. The picture vocabulary test protocol requires assessors to present a series of pictures to the students and to ask students to select the picture that corresponds to each vocabulary word spoken by the assessor. In this case, the words would have been signed; however, due to the iconic nature of many of the MSL signs, participants in the adaptation workshop felt that students could easily guess the correct picture from the sign. As a result of this feedback, the test was changed to an EVT in which students were presented with a series of isolated images and asked to provide the corresponding MSL sign. The final items and prompting questions for the EVT were validated by workshop participants who are deaf or hard of hearing.

- **Teacher expressive vocabulary test (EVT):** During the adaptation workshop, it was decided that teachers should also be asked to undergo an expressive vocabulary test to measure their MSL knowledge. The protocols for administering the teachers' EVT were the same as the student version, but the images and vocabulary were different. The items and prompting questions for the expressive vocabulary test for teachers were also validated by workshop participants who are deaf or hard of hearing.

- **Student questionnaire:** The student questionnaire asks students about contextual factors that may impact their EGRSLA scores, such as: their exposure to MSL, their hearing status, and their access to reading materials and reading support.

- **Teacher questionnaire:** The teacher questionnaire asks teachers about their preparedness to teach deaf or hard of hearing students, teaching practices, MSL-fluency level, and access to materials for teaching students who are deaf or hard of hearing.

- **Classroom observation instrument**: The classroom observation protocol was sourced from the methodology used in the Snapshot of School Management Effectiveness (SSME) battery of instruments. A SSME teacher observation is used to collect such data as pedagogical approaches used, time on task, and interactions among students. This tool can allow projects to monitor software use and its impact on teaching approaches and student learning. The contents of the SSME teacher observation form were reviewed by the adaptation team. Several adjustments were made to reflect pedagogical practices, and didactic materials specific to a deaf education context were added to the tool.

Pretest

During the adaptation workshop, assessors pretested instruments in two non-intervention schools. STS provided parameters for the selection of pretest schools including that students must be in Grades 1 or 2, in similar learning environments based on the schools' approach to language teaching and usage in the classroom, and would not receive the intervention. Initially, only one school was included in the pretest. However, after data collection, the adaptation team recognized that the use of sign language was so low that it would be very difficult to determine the tests' validity if just the initial school was used. In response, USAID identified an additional school that uses MSL as the language of instruction almost exclusively. The instruments were then pretested with students from that school as well.

As a result of the pretest, the following changes were made to the instruments:

1. The untimed diacritic name identification subtask was removed, and the timed syllable identification subtask was revised by (1) reducing the number of items from 100 letter-diacritic pairs to 25 pairs and (2) changing protocol to administer the subtask as untimed. In MSA, each diacritic is taught not as an independent mark, but in combination with the letters of the Arabic alphabet. Additionally, there is not one standard system in MSL for fingerspelling diacritic markings. In some regions, letters and diacritics are signed with the same hand, while in other regions, they are signed with two hands simultaneously. Administering syllable identification as an untimed subtask allows assessors to record whether students are using the diacritic system in the IDRT software or a regional variation.

2. The unsigned reading and reading comprehension subtask was removed from the EGRSLA. Analysis of the data revealed that students could guess some of the answers of the comprehension questions based on their general knowledge.

3. The number of items in the student EVT was reduced from 38 to 25. On the 13 items removed, there was low variability in students' responses, meaning all students or no students got the item correct.

IV. Pilot Field Preparation and Data Collection

Sample

The pilot test sample consisted of 155 students in Grades 1 and 2 at eight schools across Morocco. Because the pilot test was initially intended to serve as the operational baseline data collection, the student sample was predetermined based on schools where IDRT will be implementing the project. All students in Grades 1 and 2 from selected schools were assessed. Table 1 presents the number of students assessed by gender and grade. The number of students per school ranged from eight to 46.

Table 1: Number of Students Assessed by Gender and Grade

Category	Number of Students
Gender	
Male	91
Female	64
Grade	
Grade 1	70
Grade 2	85

Assessor Training

Assessor trainings took place October 12 and 13, 2016, prior to the pretest, and October 17 and 18, 2016, prior to the pilot. A final day of refresher training will occur on October 21, 2017 before the operational data collection begins. Given the rareness of individuals with fluency in both reading MSA and using MSL, the pool of potential assessors for this study is extremely limited.

IDRT recruited two people to act as the main assessors:
1. One who is a hearing-native user of Moroccan Arabic, reads and speaks MSA fluently, and a second-language user of MSL.
2. One who is a hard of hearing-native user of MSL, a second-language user of Moroccan Arabic, and reads and speaks MSA fluently.

Additionally, a supervisor, who is a hearing-native user of Moroccan Arabic, reads and speaks MSA, and a second-language user of MSL, was hired to conduct the teacher surveys and classroom observations. All three individuals were selected during the adaptation workshop and were present during the instrument selection and design process. The assessors all had experience working in schools that serve students who are deaf or hard of hearing.

During the pretest training, it was challenging for the assessors to reach consensus on the language to be used in instructing the students. As a result, additional time was given to the assessors to practice instructions and find agreement about how best to deliver the assessments. Many techniques were used including asking adults to simulate students' responses. During the pretest, all assessments were videotaped to provide ample feedback about what worked, scoring accuracy using the tablet, and the assessors' approach.

During the assessor training, all three participants
- reviewed EGRSLA principles and gained a comprehensive understanding of the EGRSLA instrument components
- learned and practiced EGRSLA administration and scoring procedures
- learned and practiced administering and scoring the other evaluation instruments
- standardized the translation of instructions from MSA to MSL for each student instrument
- standardized the acceptable responses for each item on each assessment subtask

- practiced conducting all components of the MSL EGRSLA on tablets and on paper
- participated in an inter-rater reliability (IRR) test of administration and scoring

Inter-rater Reliability Test

As a part of assessor training, IRR tests were conducted to calculate the consistency of assessors' agreement in their assessment decisions. IRR tests ensure that the assessors interpret student responses in the same way. Assessor agreement must meet or exceed a threshold of 90 percent for results to be considered reliable. An IRR test was done during the first day of pretest data collection, and assessor agreement did not meet the 90 percent threshold. The two assessors came to consensus on what should be the correct answer for each item, but they did not consistently mark those answers in the same way. To accommodate this, following the pretest, each assessor was given a different portion of the test to conduct with the students. This ensured reliability because most EGRSLA subtasks were conducted by one of the assessors and all vocabulary and listening comprehension subtasks were conducted by the other assessor. Assessors continued to review each other and provided feedback if any disagreements arose about the scoring. Because each assessor was responsible for a different portion of the assessment during pilot data collection, no further IRR tests were conducted.

Data Collection

Following the pretest data collection and assessor training, the two assessors and supervisor traveled to the eight schools to conduct the pilot data collection. An STS electronic data capture specialist accompanied the assessment team for the first week of data collection. One assessor administered the bulk of the EGRSLA subtasks, while the second assessor administered the MSL comprehension EGRSLA subtask, the student EVT, and the student survey. The supervisor conducted classroom observations and administered the other instruments to the teachers. The pilot data collection took place between October 25 and November 14, 2016.

Student responses to the student questionnaire and most EGRSLA subtasks—except the MSL comprehension subtask—were recorded on a tablet using the electronic data capture application Tangerine.[4] Because students had the highest fluency on the letter name identification subtask—and as a result, were often able to sign faster than an assessor could score on the tablet or paper—responses to this subtask were filmed and scored in Tangerine using video playback at the end of each day of data collection. Student responses to the EGRSLA MSL comprehension subtask and the student EVT were recorded on paper during the assessment and entered in Tangerine at the end of each day. This protocol was selected due to two factors that made it impractical for the assessors to record students' scores on the tablet in real-time: (1) only two tablets were available to assessors during data collection, and (2) videos for the MSL comprehension subtask and images for the student EVT were presented to students using the tablet. The teacher survey, teacher CDI, student CDI and teacher EVT were also all administered on paper and entered in Tangerine at the end of each day. The classroom observation was

---

[4] Tangerine® (http://tangerinecentral.org/) is an electronic data collection software designed for use on mobile computers, including netbooks, tablet computers and smartphones. Its primary use is to enable recording of children's responses in oral early grade reading and mathematics skills assessments, specifically Early Grade Reading Assessment (EGRA) and Early Grade Mathematics Assessment (EGMA), and interview responses from children, teachers and principals on home and school context information.

administered on paper and entered in Excel following data collection.

V.   Lessons Learned for Data Collection Process and Procedures

The purpose of the pilot study was to determine the reliability and validity of the EGRSLA and other language assessments for this student population as well as determine how to effectively collect high-quality, accurate data in an efficient manner. Both the pretest and the pilot data collections yielded lessons on ways to enable data collection and ensure data reliability.

The following are challenges and solutions to the issues faced:
1. Tablets, Tangerine, and paper-based tools: During the pretest, the assessors collected the data on paper. However, the assessment team felt they could not also manage tablets while looking at the student to determine if the sign was correct. Following the pretest, the assessment team recommended switching to tablets for data collection because it eased the burden of timing, tracking student progress, and marking students' responses. The responses for the EGRSLA subtasks were recorded on tablets through Tangerine while the responses for the student questionnaire and the Expressive Vocabulary Test were recorded on paper and input into Tangerine each night by the assessors.
2. Recording: Scoring fluency tasks for quick readers proved to be challenging. During the pretest, the assessors recorded all subtasks and the assessment team reviewed students' answers to determine if the answers were correct and if there was assessor agreement on the way the items were scored. Following this review, it was determined that only the letter name identification subtask would need that level of review moving forward: some students signed faster than an assessor could score. The students did not progress through other timed subtasks—familiar word reading and the reading passage—quickly enough to require a secondary check on the scoring. The assessors agreed on this method and used it during piloting. If a student read quickly, they would review the video of that student later that day to verify the scores in a second Tangerine entry.
3. Length of the assessment: During the pretest, the assessments took about 60 minutes per student. As the different subtasks were allocated to different assessors based on the challenge of securing assessor agreement, it was determined that students would spend approximately 30 minutes with one assessor, take a short break, and then complete the remainder of the assessment with the other assessor. While students were engaged, and appeared to enjoy participating in the assessment, there is an issue of test fatigue on any assessment this long. The assessors encouraged students to take a break if they needed to during the assessment.

VI.   Data Management

During the pilot, data for the EGRSLA, the additional language assessment, teacher observations, and student and teacher interviews were handled as follows:
1. Assessors collected the student survey, EVT, and EGRSLA subtask on tablets using Tangerine while in the schools.
2. The letter name identification subtask was video recorded during the assessment and the

assessor started each recording by saying the student's EGRSLA ID.
3. The supervisor conducted a teacher survey and classroom observation on paper.
4. At the end of each day, the assessors uploaded their Tangerine files to the server.
5. The assessors determined which students moved quickly enough through the letter name identification subtask to merit an additional review. The assessors created an additional record for that student, reviewed the video recording, and re-scored the letter name identification subtask. This was treated as the official file.
6. The assessors inputted the teacher observation and teacher survey data into Tangerine.
7. Once the data were collected from all students in the pilot, STS reviewed the data, merged duplicate files, cleaned all data, and translated all instruments.
8. STS conducted data analysis to determine item fit and the overall validity of the assessments.

VII.   Pilot Data Analysis and Results

When examining the quality of the EGRSLA for students who are deaf or hard of hearing in Morocco, STS examined how well the items functioned within each subtask as well as the relationships between subtasks. In an EGRA for hearing students, the EGRA directly assesses students' monolingual ability—where text, reception and production tasks occur in a single language. With students tested in MSL, the text is read in MSA while the students respond in MSL. The bilingual nature of this assessment warrants the pilot analyses to further test item function and whether assumptions of subtask function and relationships hold. For example, do students who perform better on a decoding subtask perform better on a reading fluency task? This assumption holds true for hearing students assessed in MSA, but researchers are uncertain whether it holds true for students who are deaf or hard of hearing tested in MSA and MSL.

Mean results for the EGRSLA and other language assessments are presented in Table 2.

Table 2: Mean Results for Assessments

Assessment	Timed/ Untimed	N	Mean	Total Items
Letter name identification	Timed	154	39.94	100
Syllable identification	Untimed	155	7.51	25
Syllable identification*	Untimed	155	12.80	25
Familiar word reading	Timed	154	3.02	50
Reading passage	Timed	154	3.29	61
Reading comprehension	Untimed	155	0.01	5
Student EVT	Untimed	152	6.46	25
Student EVT*	Untimed	152	15.73	25
MSL comprehension	Untimed	152	0.03	5
MSL comprehension*	Untimed	152	0.05	5

*An asterisk represents the subtask mean when regional variations of signs were considered correct in addition to the signs from the IDRT software.

The following steps were taken to analyze pilot data:

1. Descriptive statistics were computed.
2. The difference in responses between Grade 1 and Grade 2 were analyzed. This calculation provides information about the quality of the tasks and operates on the assumption that the majority of advanced students should receive better results. For example, it is expected that there will be a higher proportion of correct responses among Grade 2 students compared to Grade 1 students.
3. An item level functioning (ILF) analysis was conducted. ILF is an analysis of the quality of each item within each subtask. ILF was examined using the Rasch Model, which maps items where individual ability summarizes overall abilities (in this case, bilingual reading ability) and then "locates" each item on the presumed underlying ability of all individuals tested. This analysis allows researchers to identify items that require students to have high reading ability to get the item correct, items that require only low ability, or items answered similarly (either wrong or right) regardless of individual ability. In the last case, items are said to be "poor discriminating items" as they do not discriminate well between students of high or low ability. In turn, this may indicate that the item is poorly constructed.

   The ILF analysis in Table 3 lists the items from the EGRSLA and other language assessments. Item level feedback was determined through the Rasch Analysis as well as feedback from language experts. Student responses were considered correct using any regional variations, not just the signs from the project software. Poor discriminate items, or items that may not sufficiently discriminate between readers with high or low abilities, need to be reviewed for construction (wording, instructions, and context of testing). High difficulty items need to be reviewed to ensure they are in fact difficult due to the item's content, not due to poor item construction. For each of the subtasks, items identified as problematic are noted along with the issue that was vetted with local experts.

Table 3: ILF Rasch Analysis

Task	Item Number	Item Level Feedback
Letter name identification	1, 3	High difficulty items
	11,14,16	Poor fit items
	9	Expect only students with high reading ability to get item correct
	4,8	Expect students with high reading ability to get items correct
Syllable identification (using IDRT software signs)	2	Poor fit item
	8	All students answered item incorrectly
	13	Expect only students with high reading ability to get item correct
	14, 9, 3	Expect even students with low reading ability to get items correct
Syllable identification	20, 25	Poor fit items
	13	Expect only students with high reading ability to get item correct
Familiar word reading	8, 17, 25, 26, 29, 31, 33, 40	All students answered items incorrectly
	21	Expect students with high reading ability to get item correct
	14, 35, 47, 48	Expect even students with low reading ability to get items correct
Reading passage	1, 9, 10, 14, 18, 19, 27, 29, 30, 32-36, 38, 43, 45, 49-51, 54, 57, 58, 60	All students answered items incorrectly
Student EVT (IDRT software signs)	12,13	Poor fit items
	25, 21, 23	Expect students with high reading ability to get item correct
	3	Expect even students with low reading ability to get item correct

4. The additional language assessments were also examined by subtask to identify relationships between students' performance on each subtask. Overall, the relationships indicate that students with higher scores on the student EVT, the most basic language subtask, did not differentiate well on all tasks except for the EGRLSA's familiar word fluency subtask. In other words, students who did much better than the average on the EVT did not also perform better than average on the syllable identification, letter name identification, and reading passage subtasks. Students who had lower than average performance on the student EVT had comparable performance on the familiar word reading subtask to students whose scores were at or slightly above the mean. Similar trends were observed for students who performed higher than average on the reading passage subtask. Specifically, students who had high performance on the reading passage (two standard deviations or higher than the mean) did not perform much differently on

other subtasks than students who had only slightly higher than average scores on the reading passage subtask.

These incongruities in the data indicate that, for students tested in the pilot, the ILF results and the relationships observed between performance on subtasks may be a result of different factors: (1) a small number of students were high performers (nine students in Grade 1 and 17 in Grade 2); (2) the low number of higher difficulty items, which gave students less opportunity to demonstrate higher ability; or (3) the bilingual nature of the assessment itself (translation from Arabic print to MSL).

The following are additional observations between subtasks and students:
- Students with higher scores on the student EVT using the signs from the IDRT software also had higher scores on the letter name identification, familiar word reading, and reading passage subtasks. This relationship between the student EVT and the letter name identification, familiar word reading, and reading passage subtasks was the same for students who had higher scores on the student EVT when regional variations were considered correct.
- Higher scores on the syllable identification subtask were associated with higher scores on the student EVT.
- Higher scores on the syllable identification subtask and on the student EVT were associated with not having a zero score on all other EGRSLA subtasks (letter name identification, familiar word reading, and reading passage).
- High performance on the student EVT using IDRT software or regional signs was a predictor of reading passage fluency.
- Letter name identification and familiar word reading scores matter regardless of how students do on the student EVT or MSL comprehension subtask. Scores on both tasks were predictors of reading passage fluencies; this was true regardless of whether they had high scores using the signs from the IDRT software or including regional variations.
- On average students in Grade 2 performed better than students in Grade 1, and girls performed better than boys.
- Students who had zero scores on the reading passage subtask also had low scores on the student EVT, regardless of whether only IDRT software signs or regional variations were also considered correct. However, there was no relation to the MSL comprehension questions, regardless of whether only IDRT software signs or regional variations were also considered correct.
- Overall, the six tasks appear to reveal two underlying reading constructs, which hold regardless of whether only IDRT software signs or regional variations were also considered correct. The two constructs are as follows:
    a. Syllable identification, letter name identification, and EVT; and
    b. Reading passage, familiar word reading, and MSL comprehension
- Lower scores on student EVT predicted students receiving zero scores on the letter name identification, syllable identification, familiar word reading, and reading passage subtasks.

However, having a zero score on MSL comprehension did not predict lower scores on these subtasks.

- Performance on the letter name identification subtask is positively related to performance on the familiar word reading and reading passage subtask, with a stronger relation to performance on the familiar word reading subtask than to the reading passage subtask.
- Having a zero score on the reading comprehension subtask was associated with lower scores on the familiar word reading subtask but not on the reading passage subtask.
- The student EVT and the MSL comprehension subtask are not correlated.

VIII.   Additional Instrument Revision

Following the pilot, it became apparent that assessment items were not all functioning at the quality needed to provide an appropriate measure of students' level of reading ability or sign language proficiency. A key issue posed during the adaptation and revision process was how much the assessment should be tied to the signs used in the project software, *MSL Clip and Create*, as opposed to also allowing regional variations to provide a more accurate assessment of the language used by the students in the different regions of Morocco. The first pilot attempted to capture these differences by providing multiple scoring options for correct responses, namely (1) the student provided correct sign from IDRT software; and (2) the student provided a correct sign but not the sign in the IDRT software (a valid regional variation). The following are lessons learned for future adaptation workshops and the finalization of the EGRSLA instruments and additional language assessments for IDRT:

1. There is a need to determine the relevancy for the student population of the content derived from this EGRSLA in Morocco. The project schools do not necessarily follow the standard curriculum as general population schools. Therefore, the standards incorporated into the pilot instrument may have been misapplied. Further exploration and verification of each school's curriculum as well as student access to MSA is needed to best determine if the EGRSLA was properly constructed.
2. The signs used in the *MSL Clip and Create* software do not necessarily represent the language used in the project's schools. While many of the project schools use variations of some of the signs in *MSL Clip and Create*, there should be further exploration into the sign variations used in different regions of Morocco and further testing on how best to consider those variations during the assessment.
3. The current measures for MSL in the EGRSLA do not function as intended and yielded little clarity on students' sign language ability through the teacher-reported CDI or the student EVT. Additional work is required to better capture students' use and knowledge of sign language apart from their reading skills in MSA.
4. The assessments used to measure teachers' MSL fluency also did not present a useful picture of their language abilities. The self-report approach yielded little differentiation among teachers, and it was not clear how reliably the assessments measured teachers' knowledge and use.

Additionally, the vocabulary test revealed that, on average, teachers only knew signs for about 65 percent of the pictures presented to them. More time should be invested to develop stronger language use and fluency measures for teachers, both to provide context around students' exposure to language as well as to be able to measure possible changes in teacher knowledge due to exposure to the IDRT software.

IX.   Conclusion

This adaptation and pilot experience highlights the complex issues surrounding language and reading assessments for students who are deaf or hard of hearing. Because Morocco uses MSL with great variance across the regions, and because schools choose their own approach to teaching and using sign language in the classroom, development of a stronger assessment tool is vital. The revised EGRSLA tool should more accurately reflect students' natural language use and proficiency in MSL as well as their reading abilities in MSA. Additionally, while reading is the ultimate goal, the language deprivation many students in Morocco face should be well documented and more accurately understood to improve the students' first language skills. Ideally, this understanding should be in place before measuring their reading skills in another language, in this case MSA. The Improving Deaf Children's Reading Through Technology in Morocco project will most likely have the greatest impact on sign language acquisition and that change should be well measured, even if the reading outcomes aren't evident in the first full year of implementation.

Next steps, including additional revisions of the EGRSLA tool, assessor training, and operational data collection in fall of 2017 and spring of 2018 will allow for a better understanding of students' abilities in MSL and MSA and a more accurate assessment of the language and reading gains associated with the Improving Deaf Children's Reading Through Technology in Morocco project.

X. Annexes

1) Letter name (Timed)

Descriptive statistics

	Frequency	Percentage (%)
Letter name auto-stop	7	4.55%

	n	Mean	SD	Range
Number of letters attempted				
Overall	154	65.9	26.6	0-100
Without zero-score	147	68.6	24.2	0-100
Number of letters correct				
Overall	154	39.1	24.4	0-91
Without zero-score	147	40.9	23.4	0-91

Rasch analysis

Item no.	Item parameter	S.E.	Infit MSQ	Infit ZSTD	Outfit MSQ	Outfit ZSTD
1	-6.46	**0.56**	1.2681	0.7613	**4.0819**	**2.4741**
2	-0.89	0.22	1.0966	0.8511	1.4345	1.4614
3	-6.17	0.51	1.3514	1.0314	**3.421**	**2.0634**
4	-3.3	0.29	0.8988	-0.5391	1.3373	0.7513
5	-2.89	0.28	1.153	0.9012	1.0633	0.2911
6	-0.98	0.22	1.0785	0.6911	1.3439	1.2213
7	-1.08	0.22	1.3543	2.6514	1.4246	1.4814
8	-3.3	0.29	1.1785	1.0312	1.7752	1.3718
9	-4.05	0.32	0.8453	-0.8392	0.304	-1.1397
10	-2.97	0.28	1.1693	0.9912	0.9139	-0.0391
11	0.98	0.22	1.3219	3.1613	**1.917**	1.7719
12	0.23	0.21	1.1774	1.9912	1.4006	1.0714
13	-1.98	0.25	0.7547	-1.6692	0.6578	-1.1893
14	-3.13	0.29	1.2081	1.1812	**1.8084**	1.5018
15	-0.6	0.21	1.0387	0.401	1.013	0.141
16	-1.91	0.25	0.6434	-2.6094	**0.4558**	**-2.2395**
17	-0.38	0.21	1.0659	0.7011	1.0589	0.2811
18	1.23	0.22	1.0449	0.471	0.7873	-0.3192
19	-1.08	0.22	1.452	3.2815	**1.7523**	**2.3718**
20	-2.37	0.26	1.2382	1.3912	1.218	0.6912
21	0.85	0.21	1.2338	2.4512	1.2409	0.6612
22	0.85	0.21	1.2339	2.4512	1.2499	0.6812
23	-0.51	0.21	1.4685	4.0915	1.4362	1.3314
24	0.89	0.21	1.2755	2.8113	1.1382	0.4511

25	1.71	0.24	1.0965	0.7811	0.9511	0.091
26	0.14	0.21	1.3924	4.0914	1.2382	0.7212
27	1.6	0.24	1.1623	1.3212	0.9947	0.171
28	1.96	0.26	1.2032	1.3712	**9.9**	**7.0799**
29	-0.65	0.21	1.6135	4.9416	**2.1507**	**3.0222**
30	1.83	0.25	1.142	1.0511	1.0747	0.3211
31	-3.66	0.31	1.3219	1.7213	**2.7058**	**2.1127**
32	-2.66	0.27	1.0424	0.301	1.1527	0.5012
33	-1.85	0.25	1.0112	0.121	0.8991	-0.2591
34	1.13	0.22	1.2402	2.3012	**1.7057**	1.4117
35	1.28	0.23	1.0316	0.341	0.834	-0.1992
36	0.98	0.22	1.2249	2.2712	**1.6303**	1.3316
37	-1.67	0.24	0.8912	-0.7291	0.9201	-0.1991
38	-3.05	0.29	1.2139	1.2112	**1.5544**	1.1616
39	-2.97	0.28	1.2283	1.2912	1.4824	1.0815
40	1.28	0.23	1.0343	0.361	0.8029	-0.2692
41	1.28	0.23	1.2694	2.4213	1.1858	0.5212
42	-1.03	0.22	1.0766	0.6611	0.9716	-0.009
43	0.62	0.21	1.1592	1.7812	0.9794	0.091
44	-0.33	0.21	1.0911	0.9611	0.9897	0.081
45	0.1	0.21	1.0858	0.9911	0.8883	-0.1891
46	-1.03	0.22	1.0824	0.7111	1.0325	0.211
47	-0.89	0.22	1.1507	1.2912	1.141	0.5811
48	0.45	0.21	1.0979	1.1411	0.82	-0.3492
49	0.62	0.21	1.0583	0.6911	0.7825	-0.4392
50	-0.03	0.21	1.1599	1.7412	1.1264	0.4611
51	-0.84	0.22	0.8471	-1.3892	0.7297	-0.9593
52	-1.56	0.24	0.7864	-1.5792	0.6729	-1.2693
53	-1.62	0.24	0.6841	-2.4293	0.5645	-1.7994
54	-0.56	0.21	0.9544	-0.419	0.9069	-0.2091
55	-0.6	0.21	0.8545	-1.4391	0.6989	-1.0093
56	-1.5	0.24	0.6586	-2.7493	**0.4557**	**-2.4695**
57	-1.34	0.23	0.6748	-2.7293	0.56	-1.8894
58	1.28	0.23	1.1365	1.2911	1.1194	0.4011
59	1.28	0.23	1.1434	1.3511	1.1083	0.3811
60	0.75	0.21	0.8768	-1.4291	0.6287	-0.8794
61	-0.93	0.22	0.8424	-1.3892	0.6721	-1.2493
62	-0.89	0.22	0.8516	-1.3291	0.6727	-1.2293
63	-0.56	0.21	0.8985	-0.9891	0.7063	-0.9593
64	1.6	0.24	1.0348	0.331	0.7906	-0.2392
65	1.13	0.22	0.8368	-1.7292	0.6088	-0.8194
66	1.18	0.22	0.8268	-1.8092	0.6041	-0.8194
67	0.32	0.21	0.7784	-2.7992	0.5613	-1.2194
68	1.71	0.24	1.0016	0.061	0.8613	-0.0891

69	0.98	0.22	0.819	-2.0292	0.5916	-0.9194
70	0.45	0.21	0.9786	-0.229	0.7275	-0.6193
71	0.1	0.21	0.8882	-1.3091	0.6759	-0.8793
72	3.79	0.47	1.0952	0.3711	0.66	-0.3893
73	-0.16	0.21	0.776	-2.6292	0.5995	-1.2294
74	0.53	0.21	0.9495	-0.5691	0.8921	-0.1391
75	1.71	0.24	1.0087	0.111	0.7715	-0.2792
76	1.03	0.22	0.9193	-0.8391	0.6137	-0.8394
77	0.8	0.21	0.948	-0.5591	0.6857	-0.6893
78	1.38	0.23	0.8027	-1.9092	0.5261	-0.9695
79	1.43	0.23	1.039	0.381	0.7082	-0.4593
80	0.49	0.21	0.8193	-2.2292	0.7107	-0.6693
81	0.89	0.21	0.6983	-3.6793	**0.4611**	-1.4095
82	0.89	0.21	0.6983	-3.6793	**0.4611**	-1.4095
83	1.08	0.22	0.7277	-3.0893	**0.4683**	-1.2995
84	1.13	0.22	0.6944	-3.4593	**0.4441**	-1.3596
85	1.08	0.22	0.7084	-3.3393	**0.4551**	-1.3495
86	1.13	0.22	0.6985	-3.4093	**0.4467**	-1.3496
87	1.33	0.23	0.7073	-3.0393	**0.4389**	-1.2796
88	1.28	0.23	0.7232	-2.9193	**0.4534**	-1.2595
89	2.03	0.26	0.7398	-1.8693	**0.4017**	-1.2296
90	1.38	0.23	0.7391	-2.6093	**0.4514**	-1.2095
91	1.54	0.24	0.8863	-0.9691	1.2779	0.6613
92	1.71	0.24	0.8713	-1.0191	0.5109	-0.9295
93	1.6	0.24	0.8366	-1.3992	0.5021	-0.9695
94	1.54	0.24	0.8024	-1.7792	**0.4847**	-1.0295
95	1.65	0.24	0.8184	-1.5292	**0.4911**	-0.9995
96	2.77	0.32	0.9383	-0.2091	0.5063	-0.8595
97	2.77	0.32	0.9101	-0.3391	**0.4938**	-0.8995
98	2.58	0.3	0.8639	-0.6391	**0.4588**	-0.9995
99	1.96	0.26	0.9518	-0.289	0.5429	-0.8095
100	2.17	0.27	0.9456	-0.2891	0.5234	-0.8395

```
--------------------------------------------------------------------------------
MEASURE    PERSON - MAP - ITEM
                     <more>|<rare>
    4                     +
                          |
                          |   i72
                          |
                          |
                          |T
    3            X  T+
                          |
                          |
                          |   i96   i97
                    X     |   i98
                          |
    2            XX  +  i10
                 XX  |   i89
              XXXXX  |S i30   i99
                          |   i25   i68   i75   i92   i95
                XXX  |   i27   i64   i91   i93   i94
              XXXX S|   i78   i79   i87   i90
    1        XXXXX  +  i18   i35   i40   i41   i58   i59   i66   i88
             XXXXX  |   i11   i34   i36   i65   i69   i76   i83   i84   i85   i86
              XXXX  |   i21   i22   i24   i81   i82
     XXXXXXXXXXXX  |   i60   i77
             XXXXX  |   i43   i49   i74   i80
             XXXXX  |   i48   i70
    0        XXXXXX  +M i12   i67
            XXXXXXX  |   i26   i45   i71
             XXXXX  |   i50   i73
             XXXXX  |   i44
               XXX  |   i17
             XXXXX M|   i15   i23   i54   i55   i63
   -1      XXXXXXXX  +
               XXX  |   i2    i47   i51   i6    i61   i62
                XX  |   i42   i46   i7
                XX  |
                          |S i57
               XXX  |   i52   i53
   -2            X  +  i37
                XX  |   i33
                 X  |   i13
                 X  |
              XXXX  |   i20
               X S|
   -3            X  +  i32
                XX  |T i5
                XX  |   i10   i38   i39
                          |
               XXX  |   i4    i8
                 X  |
   -4        XXXXX  +
                          |
             XXXXXX  |
                          |   i9
              XXXX  |
                          T|
   -5            +
               XXX  |
                          |
                          |
                          |
   -6 XXXXXXXXXXXX  +
                     <less>|<freq>


TABLE 1.3 letter_names.dta                    ZOU421WS.TXT   Feb  6 2017 17:14
INPUT: 154 PERSON  100 ITEM   REPORTED: 154 PERSON  91 ITEM  2 CATS WINSTEPS 3.91.0
--------------------------------------------------------------------------------
```

2) Syllable identification

Descriptive statistics

	n	Mean	SD	Range
Number of items correct: correct only				
Overall	154	7.6	5.2	0-17
Without zero-score	121	9.6	3.8	1-17
Number of items correct: correct but different sign				
Overall	154	12.9	8.4	0-25
Without zero-score	121	16.4	5.7	1-25

Rasch analysis – IDRT software signs only

Item no.	Item parameter	S.E.	Infit MSQ	Infit ZSTD	Outfit MSQ	Outfit ZSTD
1	-3.57	0.35	1.0549	0.3011	0.8632	-0.0791
2	2.13	0.32	1.0661	0.3711	**2.336**	**2.1323**
3	2.61	0.38	0.9575	-0.049	0.4725	-0.9195
4	0.45	0.22	0.9086	-0.9491	1.5696	1.7716
5	-3.04	0.3	1.1438	0.7611	1.0976	0.3711
6	-0.46	0.21	1.2377	2.5712	1.2444	1.2612
7	0.65	0.23	1.1009	0.9911	1.0447	0.241
8	6.04	**1.83**	1	0	1	0
9	2.77	0.4	0.9639	-0.009	0.5754	-0.5894
10	0.6	0.22	1.1175	1.1611	1.0002	0.111
11	-2.7	0.28	0.8299	-0.9992	0.6608	-1.0193
12	-0.29	0.21	1.1409	1.5911	1.2032	1.0112
13	-5.08	0.56	1.108	0.3911	0.941	0.2609
14	3.14	0.47	0.9817	0.081	0.62	-0.3694
15	-0.11	0.21	1.1128	1.2811	0.9941	0.051
16	-1.1	0.22	0.7384	-2.9193	0.6491	-2.2694
17	1.94	0.3	0.9277	-0.3091	1.5291	1.1315
18	0.75	0.23	0.9549	-0.389	0.8735	-0.2991
19	-2.63	0.27	0.7728	-1.4392	0.8099	-0.5092
20	-0.2	0.21	1.1614	1.8012	1.2471	1.1612
21	1.47	0.26	1.1834	1.2112	1.5568	1.3016
22	-1.24	0.22	0.6559	-3.8393	0.5343	-3.1495
23	0.7	0.23	0.8328	-1.6592	0.7161	-0.8793

24	1.61	0.27	0.9679	-0.149	0.7194	-0.5693
25	1.61	0.27	1.0235	0.201	0.9452	0.0209

```
MEASURE     PERSON - MAP - ITEM
                 <more>|<rare>
   4               +
                   |
                   |
                   |
                   |   i14
                   |
   3               +
                   |   i9
                   |   i3
                   |
                 T |
                   |S
   2               +   i17
              #    |
                   |   i24      i25
              #    |   i21
                   |
            .##    |
   1        ###    +
                 S |   i18      i23
             .#    |   i10      i7
                   |   i4
           ####    |
                   |
   0       .###   +M  i15
           .###    |   i12      i20
                   |   i6
          .#### M |
                   |
           ####    |
  -1               +   i16
                   |   i22
            .##    |
                   |
            .##    |
                   |
  -2             S +
           .###   |S
                   |
                   |   i19
             .#    |   i11
                   |
  -3               +   i5
                   |
             .  T |
                   |   i1
                   |
                   |
  -4         #     +
                   |
                   |T
                   |
                   |
                   |
  -5         .     +   i13
                   |
                   |
                   |
                   |
                   |
  -6   ###########  +
                 <less>|<frequent>
EACH "#" IS 3: EACH "." IS 1 TO 2
```

Rasch analysis – including regional variations

Item no.	Item parameter	S.E.	Infit MSQ	Infit ZSTD	Outfit MSQ	Outfit ZSTD
1	-2.21	0.38	1.2514	1.0313	1.0227	0.281
2	-1.59	0.33	0.9394	-0.2291	0.5102	-0.8495
3	2.3	0.23	1.06	0.6111	1.0363	0.221
4	-0.93	0.29	0.8868	-0.6091	1.1	0.3711
5	-1.82	0.35	1.3487	1.5213	1.1677	0.4612
6	0.11	0.24	0.9109	-0.6491	0.7376	-1.0293
7	2.46	0.23	1.1504	1.3912	1.1535	0.5412
8	1.56	0.22	0.9144	-0.9091	0.7764	-0.8392
9	1.07	0.22	1.0334	0.361	0.9827	0.011
10	2.14	0.22	1.1038	1.0611	1.0177	0.161
11	-1.38	0.31	0.6042	-2.3094	0.291	-1.7697
12	-1.82	0.35	0.8472	-0.6692	0.5819	-0.5794
13	-4.57	0.78	0.805	-0.1392	0.225	-0.8198
14	1.36	0.22	1.1139	1.1911	1.2263	0.9112
15	1.02	0.22	0.9086	-0.8991	0.753	-1.0392
16	0.45	0.23	0.8107	-1.6692	0.6684	-1.4993
17	0.61	0.23	0.9803	-0.139	0.8793	-0.4591
18	2.62	0.24	0.9724	-0.209	1.0388	0.231
19	-1.29	0.31	0.7569	-1.3292	0.5019	-1.0795
20	0.11	0.24	**1.5969**	**3.8916**	1.7109	2.3617
21	-1.49	0.32	1.0277	0.201	1.8078	1.3818
22	0.11	0.24	0.8175	-1.4392	0.6261	-1.5894
23	-0.69	0.28	0.8093	-1.1792	0.555	-1.3394
24	0.92	0.22	0.7991	-2.0492	0.6336	-1.6994
25	0.97	0.22	1.3172	2.8613	**2.0337**	**3.422**

```
MEASURE    PERSON - MAP - ITEM
                 <more>|<rare>
    5          .  +
                 T|
                  |
                  |
              ##  |
                  |
    4             +
                  |T
                  |
           .###   |
                  |
                  |
    3          S+
           .###   |    i18
                  |    i7
              ##  |    i3
                  |    i10
                  |
    2       .#### +
                  |S
           .####  |    i8
                  |    i14
           .###   |
           .## M|     i15      i9
    1             +     i24
              .   |
                  |    i17
              ##  |    i16
                  |
            .##   |    i22      i6
    0       .#   +M
                  |
              #   |
                  |
            # S|     i23
                  |
   -1         .  +     i4
                  |
              .   |    i19
                  |    i11      i21
              #   |    i2
                  |S i12      i5
   -2        .#  +
                  |
              .   |    i1
                 T|
                  |
              .   |
   -3             +
                  |
                  |
              .   |
                  |T
   -4             +
                  |
                  |
                  |
              .   |    i13
                  |
   -5    ########## +
                 <less>|<frequent>
EACH "#" IS 3: EACH "." IS 1 TO 2
```

3) Familiar word reading (Timed)

Descriptive statistics

	Frequency	Percentage (%)
Letter name auto-stop	107	69.5%

	n	Mean	SD	Range
Number of words attemped				
Overall	154	14.5	15.2	5-50
Without zero-score	47	35.9	9.3	16-50
Number of words correct				
Overall	154	2.9	5.6	0-36
Without zero-score	47	9.6	6.3	2-36

Rasch analysis

Item no.	Item parameter	S.E.	Infit MSQ	Infit ZSTD	Outfit MSQ	Outfit ZSTD
1	-1.66	0.34	**1.55**	**3.3916**	**1.7065**	**2.0717**
2	-0.01	0.43	1.2076	0.8412	**2.8698**	**2.3229**
3	-3.99	0.42	1.2763	1.1813	**2.4881**	**2.2225**
4	0.67	0.53	1.0527	0.2611	1.0975	0.4011
5	0.19	0.46	0.9114	-0.2191	1.2611	0.5913
6	1.94	0.82	1.5156	0.8815	1.8707	0.9719
7	-1.06	0.36	1.2889	1.7913	1.6706	1.7117
8	4.26	1.9	**1**	**0**	**1**	**0**
9	-1.43	0.35	1.2692	1.8013	1.23	0.8012
10	-0.01	0.43	1.0357	0.221	0.7509	-0.2592
11	-3.23	0.37	1.3031	1.5613	1.449	1.0014
12	-0.36	0.4	0.8153	-0.8592	0.563	-0.8894
13	1.38	0.68	1.4201	0.8714	0.9473	0.3409
14	2.85	1.12	0.4091	-0.5796	0.0391	-0.93
15	-0.36	0.4	0.8879	-0.4791	0.6263	-0.7094
16	-0.51	0.39	1.2996	1.4813	1.2935	0.7413
17	4.26	1.9	**1**	**0**	**1**	**0**
18	0.99	0.59	0.8422	-0.2192	0.5264	-0.2395
19	-0.36	0.4	0.8949	-0.4491	0.6924	-0.5293
20	0.19	0.46	1.3513	1.1914	1.2833	0.6113
21	-4.58	0.48	0.7704	-0.7592	0.7663	-0.1592
22	0.99	0.59	0.768	-0.4092	0.3569	-0.5296
23	1.38	0.68	0.8336	-0.1392	0.395	-0.2996
24	-1.66	0.34	0.8704	-0.9091	0.9929	0.071
25	4.26	1.9	**1**	**0**	**1**	**0**

26	4.26	1.9	**1**	**0**	**1**	**0**
27	-1.06	0.36	1.2719	1.7013	1.1592	0.5512
28	-3.51	0.38	0.7516	-1.3092	0.6457	-0.7094
29	4.26	1.9	**1**	**0**	**1**	**0**
30	-2.97	0.36	0.7891	-1.2492	0.6684	-0.7293
31	4.26	1.9	**1**	**0**	**1**	**0**
32	-0.01	0.43	0.9028	-0.2991	0.5831	-0.6294
33	4.26	1.9	**1**	**0**	**1**	**0**
34	-1.06	0.36	0.8017	-1.3492	0.602	-1.2094
35	2.85	1.12	1.4839	0.7715	1.3573	0.7014
36	1.94	0.82	1.49	0.8615	0.8674	0.3409
37	-0.01	0.43	0.9204	-0.2291	0.5659	-0.6694
38	-0.01	0.43	0.8025	-0.7392	0.8149	-0.1292
39	-0.8	0.37	0.6803	-2.0893	0.48	-1.5095
40	4.26	1.9	**1**	**0**	**1**	**0**
41	-0.01	0.43	0.8799	-0.3991	0.5487	-0.7095
42	0.99	0.59	0.9314	-0.0091	0.5159	-0.2595
43	0.67	0.53	0.782	-0.4892	0.3885	-0.6696
44	0.99	0.59	0.8778	-0.1391	0.4245	-0.4096
45	0.67	0.53	0.7684	-0.5392	0.3716	-0.6996
46	0.99	0.59	0.8778	-0.1391	0.4245	-0.4096
47	2.85	1.12	0.4091	-0.5796	0.0391	-0.93
48	2.85	1.12	1.4676	0.7615	0.9636	0.451
49	1.94	0.82	0.5931	-0.4794	0.151	-0.6198
50	1.38	0.68	0.7184	-0.3993	0.2993	-0.4697

```
MEASURE      PERSON - MAP - ITEM
                  <more>|<rare>
    3               +
                    |   i14     i35     i47     i48
                .   |
                    |
                    |
                    |
    2               +   i36     i49     i6
                   |S
                    |
                    |
                    |   i13     i23     i50
                    |
    1             T+   i18     i22     i42     i44     i46
                    |
                    |   i4      i43     i45
                    |
                .   |
                    |   i20     i5
    0           .  +M i10     i32     i37     i38     i41
                    |
                    |   i12     i15     i19
                .  S|   i16
                .   |
                .   |   i39
   -1           .   +
                .   |   i27     i34     i7
                .   |
                    |   i9
                .   |   i24
                   |S
   -2           .   +
                  M|
                .   |
                .   |
   -3           .   +
                    |   i30
                    |
                .   |   i11
                    |   i28
                  S|T
   -4           .   +
                    |   i3
                    |
                    |
                .   |
                    |   i21
   -5               +
                    |
                  T|
                .   |
                    |
                    |
   -6   .##########  +
                  <less>|<frequent>
EACH "#" IS 9: EACH "." IS 1 TO 8
```

4) Reading passage (Timed)

Descriptive statistics

	Frequency	Percentage (%)
Letter name auto-stop	93	60.4%

	n	Mean	SD	Range
Number of words attemped				
Overall	154	15.8	15.8	5-61
Without zero-score	61	32.3	13.6	12-61
Number of words correct				
Overall	154	3.2	4.7	0 -22
Without zero-score	61	8	4.2	1-22

Rasch analysis

Item no.	Item parameter	S.E.	Infit MSQ	Infit ZSTD	Outfit MSQ	Outfit ZSTD
1	3.71	1.83	**1**	**0**	**1**	**0**
2	0.9	0.55	1.2065	0.6112	2.4867	1.4525
3	-3.15	0.31	**2.0959**	**5.9721**	**3.6391**	**6.2336**
4	-1.18	0.33	1.0155	0.151	0.927	-0.0691
5	-2.96	0.31	1.497	3.1915	**1.9776**	**3.202**
6	-3.65	0.33	0.7417	-1.7293	0.8956	-0.2091
7	-0.72	0.35	1.0789	0.4811	0.804	-0.2692
8	2.46	1.03	1.0709	0.3811	0.7517	0.2608
9	3.71	1.83	**1**	**0**	**1**	**0**
10	3.71	1.83	**1**	**0**	**1**	**0**
11	2.46	1.03	1.0709	0.3811	0.7517	0.2608
12	1.24	0.62	0.9202	-0.0091	0.7419	0.0707
13	-6.26	0.63	0.9548	0.071	**7.3878**	**2.9174**
14	3.71	1.83	**1**	**0**	**1**	**0**
15	-2.05	0.3	1.0515	0.4311	1.0109	0.131
16	-4.23	0.36	0.8464	-0.7692	0.6114	-0.8394
17	2.46	1.03	1.0709	0.3811	0.7517	0.2608
18	3.71	1.83	**1**	**0**	**1**	**0**
19	3.71	1.83	**1**	**0**	**1**	**0**
20	-1.28	0.32	1.3618	2.1914	1.2952	0.8813
21	-3.87	0.34	0.7134	-1.8193	0.4949	-1.5495
22	2.46	1.03	1.0014	0.311	0.2994	-0.2597
23	-1.07	0.33	0.9489	-0.2691	0.7143	-0.6593
24	-0.31	0.38	0.817	-0.8092	0.5701	-0.6794

25	-3.05	0.31	0.6239	-3.0794	0.4996	-2.3595
26	0.62	0.5	1.0546	0.2711	1.1962	0.5012
27	3.71	1.83	**1**	**0**	**1**	**0**
28	1.24	0.62	1.0297	0.221	0.3992	-0.4096
29	3.71	1.83	**1**	**0**	**1**	**0**
30	3.71	1.83	**1**	**0**	**1**	**0**
31	-0.46	0.37	1.0684	0.3911	0.8493	-0.0992
32	3.71	1.83	**1**	**0**	**1**	**0**
33	3.71	1.83	**1**	**0**	**1**	**0**
34	3.71	1.83	**1**	**0**	**1**	**0**
35	3.71	1.83	**1**	**0**	**1**	**0**
36	3.71	1.83	**1**	**0**	**1**	**0**
37	0.01	0.42	0.9164	-0.2491	0.4833	-0.7595
38	3.71	1.83	**1**	**0**	**1**	**0**
39	2.46	1.03	0.7288	-0.0193	0.0866	-0.7599
40	1.7	0.74	1.0005	0.211	0.3504	-0.3196
41	-0.16	0.4	0.8045	-0.7992	0.4987	-0.7995
42	0.39	0.46	0.9291	-0.1291	0.4257	-0.6896
43	3.71	1.83	**1**	**0**	**1**	**0**
44	1.24	0.62	0.965	0.091	0.3337	-0.5397
45	3.71	1.83	**1**	**0**	**1**	**0**
46	0.9	0.55	0.6954	-0.6993	0.2345	-0.9198
47	1.7	0.74	1.0253	0.251	0.3762	-0.2796
48	2.46	1.03	0.7288	-0.0193	0.0866	-0.7599
49	3.71	1.83	**1**	**0**	**1**	**0**
50	3.71	1.83	**1**	**0**	**1**	**0**
51	3.71	1.83	**1**	**0**	**1**	**0**
52	2.46	1.03	0.7288	-0.0193	0.0866	-0.7599
53	0.19	0.44	0.6164	-1.4794	0.2754	-1.2397
54	3.71	1.83	**1**	**0**	**1**	**0**
55	2.46	1.03	1.0014	0.311	0.2994	-0.2597
56	0.9	0.55	0.9402	-0.0091	0.3397	-0.6597
57	3.71	1.83	**1**	**0**	**1**	**0**
58	3.71	1.83	**1**	**0**	**1**	**0**
59	1.24	0.62	0.7812	-0.3192	0.2788	-0.6497
60	3.71	1.83	**1**	**0**	**1**	**0**
61	2.46	1.03	0.7288	-0.0193	0.0866	-0.7599

5) EVT (Untimed)

Descriptive statistics

	n	Mean	SD	Range
Number of items correct: correct only				
Overall	152	6.5	3.9	0-16
Without zero-score	142	3.4	4.8	0-22
Number of items correct: correct but different sign				
Overall	152	15.7	6.2	0-25
Without zero-score	152	3.2	4.7	0-22

Rasch analysis – IDRT software signs only

Item no.	Item parameter	S.E.	Infit MSQ	Infit ZSTD	Outfit MSQ	Outfit ZSTD
1	-0.07	0.22	1.1141	1.0411	1.1913	0.7812
2	2.23	0.46	1.0504	0.2611	1.8599	1.2019
3	3.19	0.72	0.9666	0.171	0.3249	-0.7597
4	0.81	0.27	0.9653	-0.139	0.9773	0.091
5	-1.52	0.19	1.0551	0.7111	1.0331	0.301
6	1.15	0.3	1.0219	0.171	0.6911	-0.5193
7	1.86	0.4	0.9135	-0.1591	0.3895	-1.0796
8	0.13	0.23	1.1925	1.5412	1.1249	0.5111
9	-0.02	0.22	1.013	0.151	1.1382	0.5811
10	-0.02	0.22	0.8839	-1.0391	0.7627	-0.8792
11	-0.76	0.2	0.7558	-3.1892	0.62	-2.4494
12	1.71	0.38	1.1065	0.4411	**2.7383**	**2.2027**
13	-2.54	0.21	1.4153	3.6714	**1.5894**	**3.2116**
14	2.47	0.52	0.9958	0.141	0.5867	-0.3894
15	1.46	0.34	1.0888	0.4211	0.9921	0.171
16	2.47	0.52	0.9958	0.141	1.2554	0.5613
17	-0.07	0.22	0.8935	-0.9691	0.7606	-0.9192
18	-1.26	0.19	0.9534	-0.579	1.1335	0.9811
19	-1.41	0.19	0.8355	-2.2092	0.7571	-1.9892
20	-0.84	0.2	1.0196	0.271	0.8894	-0.6291
21	-2.86	0.22	0.9411	-0.4991	0.9665	-0.119
22	-0.16	0.22	1.0322	0.341	0.9781	-0.009
23	-2.67	0.21	0.7982	-2.0192	0.707	-1.8693
24	0.03	0.22	1.0311	0.311	0.927	-0.1791
25	-3.31	0.23	1.0323	0.291	0.7535	-1.0292

```
MEASURE      PERSON - MAP - ITEM
                  <more>|<rare>
      4              +
                     |
                     |
                     |T
                     |
                     |   i3
      3              +
                     |
                     |
                     |   i14     i16
                     |
                     |   i2
      2              +
                     |S  i7
                     |
                     |   i15
                     |
                     |   i6
      1          .  T+
                     |   i4
                     |
            ##       |
                     |
            ####     |   i8
      0              +M  i1      i10     i17     i24     i9
            .##      |   i22
          .##### S|
                     |
            .####    |
                     |   i11     i20
     -1     #######  +
                     |
            ######   |   i18
                     |   i19
          ####### M|    i5
                     |S
     -2     ######   +
                     |
            .######  |
                     |
                     |
            .####    |   i23
     -3            S+   i21
                     |
                     |
            .###### |T i25
                     |
     -4              +
                     |
            ##### T|
                     |
                     |
                     |
     -5    .######## +
                  <less>|<frequent>
EACH "#" IS 2: EACH "." IS 1
```

Rasch analysis – including regional variations

Item no.	Item parameter	S.E.	Infit MSQ	Infit ZSTD	Outfit MSQ	Outfit ZSTD
1	0.72	0.19	1.3142	3.3813	1.4124	2.0214
2	1.24	0.19	1.0954	1.1811	1.0456	0.281
3	1.93	0.2	1.1259	1.4311	1.1327	0.5111
4	0.29	0.2	1.1013	1.0711	1.0188	0.161
5	0.83	0.19	0.8827	-1.4291	0.7877	-1.1292
6	0.45	0.2	1.0124	0.171	0.9206	-0.3891
7	-0.17	0.21	1.0609	0.5811	1.0724	0.4211
8	-0.93	0.24	0.9315	-0.4491	0.6832	-1.1593
9	0.49	0.2	1.0412	0.481	0.9698	-0.109
10	1.05	0.19	0.9941	-0.049	0.9338	-0.2591
11	0.45	0.2	0.8584	-1.6091	0.776	-1.2792
12	-0.5	0.22	0.9722	-0.189	0.8133	-0.7792
13	-2.22	0.33	1.0501	0.2911	0.9979	0.181
14	0.68	0.19	1.094	1.0911	1.0784	0.4711
15	0.87	0.19	0.9401	-0.7091	0.8156	-0.9492
16	1.2	0.19	1.0147	0.211	0.9268	-0.2691
17	1.35	0.19	1.0184	0.251	0.9179	-0.2791
18	-0.55	0.22	0.8818	-0.9591	0.9762	-0.019
19	-0.4	0.22	0.7637	-2.1492	0.9234	-0.2791
20	-0.87	0.24	1.0492	0.401	0.8238	-0.5892
21	-0.31	0.22	0.9024	-0.8491	0.8315	-0.7692
22	-2.7	0.37	1.3894	1.3914	1.2128	0.5212
23	-0.93	0.24	0.7534	-1.8692	0.6292	-1.4094
24	-0.13	0.21	1.0976	0.9211	1.3892	1.8114
25	-1.83	0.29	0.8958	-0.4891	0.5038	-1.1895

```
MEASURE     PERSON - MAP - ITEM
                <more>|<rare>
    4          .  +
                T|
                 |
              ##  |
                 |
                 |
    3           +
            .### |
                 |
                 |
          ###### |
              S|T
                 |
    2     ########  +
                 |   i3
            .#### |
                 |
             .### |
                 |    i17    i2
          ###### |S i16
    1       ###  +   i10
              M|   i15    i5
            ##### |   i1     i14
              ### |
                 |   i11    i6     i9
          .#### |   i4
             .## |
    0        .### +M
                 |   i24    i7
              ## |   i21
               # |   i19
                 |   i12    i18
             .# S|
               # |   i20    i23    i8
   -1           +
             .## |S
                 |
             .## |
                 |
              .# |
                 |   i25
   -2           +
                 |
              T|T i13
               # |
                 |
                 |   i22
                 |
   -3         .  +
                 |
                 |
                 |
                 |
              ## |
   -4         ## +
             <less>|<frequent>
EACH "#" IS 2: EACH "." IS 1
```

Regression: EVT and MSL comprehension- including regional variations

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.167[a]	.028	.015	5.12302
2	.729[b]	.532	.513	3.60356

a. Predictors: (Constant), G1, male

b. Predictors: (Constant), G1, male, MSLcorrect_anySign, EVTcorrect_anySign, word_per_2minute, letter_per_2minute

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	112.773	2	56.387	2.148	.120[b]
	Residual	3910.552	149	26.245		
	Total	4023.325	151			
2	Regression	2140.412	6	356.735	27.472	.000[c]
	Residual	1882.914	145	12.986		
	Total	4023.325	151			

a. Dependent Variable: orf_per_2minute

b. Predictors: (Constant), G1, male

c. Predictors: (Constant), G1, male, MSLcorrect_anySign, EVTcorrect_anySign, word_per_2minute, letter_per_2minute

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	4.583	.752		6.095	.000
	male	-1.299	.846	-.124	-1.536	.127
	G1	-1.157	.835	-.112	-1.386	.168
2	(Constant)	-1.868	1.037		-1.801	.074
	male	.324	.610	.031	.531	.596
	G1	1.235	.636	.120	1.941	.054
	EVTcorrect_anySign	.107	.061	.128	1.740	.084
	MSLcorrect_anySign	-.855	1.423	-.035	-.601	.549
	letter_per_2minute	.027	.015	.134	1.750	.082
	word_per_2minute	.550	.055	.640	9.984	.000

a. Dependent Variable: orf_per_2minute

Regression : EVT and MSL comprehension– IDRT software signs only

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.167[a]	.028	.015	5.12302
2	.729[b]	.531	.512	3.60610

a. Predictors: (Constant), G1, male

b. Predictors: (Constant), G1, male, MSLcorrect_inProgSign, EVTcorrect_inProgSign, word_per_2minute, letter_per_2minute

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	112.773	2	56.387	2.148	.120[b]
	Residual	3910.552	149	26.245		
	Total	4023.325	151			
2	Regression	2137.755	6	356.293	27.399	.000[c]
	Residual	1885.570	145	13.004		
	Total	4023.325	151			

a. Dependent Variable: orf_per_2minute

b. Predictors: (Constant), G1, male

c. Predictors: (Constant), G1, male, MSLcorrect_inProgSign, EVTcorrect_inProgSign, word_per_2minute, letter_per_2minute

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	4.583	.752		6.095	.000
	male	-1.299	.846	-.124	-1.536	.127
	G1	-1.157	.835	-.112	-1.386	.168
2	(Constant)	-.782	.887		-.882	.379
	male	.278	.610	.027	.456	.649
	G1	1.245	.637	.121	1.956	.052
	EVTcorrect_inProgSign	-.018	.095	-.014	-.191	.849
	MSLcorrect_inProgSign	-3.254	1.854	-.101	-1.755	.081
	letter_per_2minute	.045	.015	.225	3.008	.003
	word_per_2minute	.567	.056	.659	10.210	.000

Factor Analysis

Component Loadings

	EVT and MSL in Any Sign		EVT and MSL in Program Sign	
	Factor 1	Factor 2	Factor 1	Factor 2
Letter	.830	.283	.915	.169
Syllable	.827	.296	.885	.220
EVT	.710	.316	.789	.257
MSL	.546	-.440	.280	.065
ORF	.233	.848	.203	.899
Word	.321	.783	.231	.891

Mean Scores by Zero-Score Status on Reading Passage

	grade			
	1		2	
	reader	nonreader	reader	nonreader
	Mean	Mean	Mean	Mean
Letter Fluency	48.74	19.71	58.06	40.54
Word Fluency	3.13	.57	7.31	1.97
Oral Reading Fluency	7.63	0.00	8.75	0.00
EVT Correct (in Program Sign)	8	4	8	7
EVT Correct (in any Sign)	19	12	19	15
MSL Correct (in program sign)	0			
MSL Correct (in any sign)				
Syllable Correct (in Program Sign)	10	3	10	8
Syllable Correct (in any Sign)	17	6	17	14