



Integrated intelligent LEARNing environment for Reading and Writing
D7.2 – Testbed Implementation



Document identifier	D7.2 Evaluation Testbed Implementation_draft.docx
Date	2015-06-30
WP	WP7
Partners	EPIRUS, DYSACT, IOE
WP Lead Partner	DYSACT
Document status	Final
Deliverable Number	D7.2
Deliverable Title	Testbed Implementation
Deliverable version number	V1.0
Work package	WP7
Task	Task 7.2 Completion of Testbed implementation
Nature of the deliverable	<u>Report (R)</u>
Dissemination level	Public (PU)
Date of Version	2015-06-30
Author(s)	D. Lukes V. Zakopoulou M. Vasalou
Contributor(s)	L. Benton D. Gooch J. Flower
Reviewer(s)	A. Symvonis

Abstract	This report describes the test-bed implementation of the iLearnRW project evaluation. It describes how the plans outlined in deliverable 7.2 were implemented, changes that took place and lessons learned. It outlines data that was collected and lists the tools that were used in their collection. This report is preliminary to the final evaluation report (deliverable 7.3) which will describe the analysis of the data and provide final conclusions.
Keywords	Evaluation, research, design, qualitative, quantitative.

Document Status Sheet

Issue	Date	Comment	Author
v01	2015-06-15	Initial outline and draft	Dominik Lukes
v02	2014-06-26	Description of data collected and incorporation of Greek data	Dominik Lukes
v03	2014-06-29		Mina Vasalou
v04	2014-06-30		Dominik Lukes

Project information

Project acronym:	ILearnRW
Project full title:	Integrated Intelligent Learning Environment for Reading and Writing
Proposal/Contract no.:	318803

Project Officer: Krister Olson

Address:	L-2920 Luxembourg, Luxembourg
Phone:	+35 2430 134 332
E-mail:	krister.olson@ec.europa.eu

Project Co-ordinator: Noel Duffy

Address:	Dolphin Computer Access Ltd. Technology House, Blackpole Estate West, Worcester, UK. WR3 8TJ
Phone:	+01 905 754 577
Fax:	+01 905 754 559
E-mail:	noel.duffy@yourdolphin.com

Table of Contents

1. SUMMARY.....	4
2. METHODOLOGY AND EVALUATION ACTIVITIES.....	5
2.1. PARTICIPANTS	5
2.1.1 <i>Design-Based Research – UK.....</i>	5
2.1.1 <i>Programme Evaluation and Case Study Research – UK.....</i>	6
2.1.1 <i>Programme Evaluation – Greece</i>	6
2.2. TECHNOLOGY.....	6
2.3. UK IMPLEMENTATION.....	7
2.3.1 <i>Design-Based Research (DBR).....</i>	7
2.3.1 <i>Programme Evaluation.....</i>	8
2.3.1 <i>Case Study Research.....</i>	11
2.4. GREEK IMPLEMENTATION OF PROGRAMME EVALUATION.....	11
2.5. SUMMARY OF DEVIATIONS FROM ORIGINAL EVALUATION PLANS	12
3. CONCLUSIONS AND LESSONS LEARNED.....	13
3.1. WORKING WITH STUDENTS	13
3.2. WORKING WITH SCHOOLS	14
3.3. HARDWARE	14
3.4. SOFTWARE DEVELOPMENT AND DEPLOYMENT	15
4. APPENDICES	17
4.1. SCREENING TEST COMPLETE DATA.....	17
4.1.1 <i>Initial screening English.....</i>	17
4.1.1 <i>Final screening English.....</i>	17
4.1.1 <i>Screening test Greek</i>	18
4.2. INTERVIEW QUESTIONNAIRES.....	18
4.2.1 <i>Half Term Student Questionnaire</i>	18
4.2.1 <i>Pre-Independent Study Interview.....</i>	19
4.2.1 <i>Final interview.....</i>	19
4.3. LIST OF DOCUMENTS DEVELOPED TO PROVIDE GUIDANCE WITH THE USE OF THE GAME, READER AND TABLET	20
4.4. MID-EVALUATION COMMUNICATION WITH PARENTS.....	21
4.5. SAMPLE OF RAW LOG DATA	23
4.6. SAMPLE OF PRE-PROCESSED LOG DATA	26
4.7. SAMPLE OF VISUALISATION OF LOG DATA	27
4.8. FINAL CERTIFICATE OF APPRECIATION	29

1. Summary

We have completed the data collection phase of the evaluation during a school-based evaluation period running from October 2014 until June 2015. The bulk of the data was collected during large-scale school trials of the iLearnRW system between January and June 2015.

In both countries, the evaluation followed a similar pattern with differences in implementation as outlined in deliverable 7.1 on evaluation plans. While there were some divergences (described below) in the timing of certain key events, we were able to:

- Conduct programme evaluation with over 60 students in each language between November 2014 and June 2015
- Conduct design-based research evaluation with 9 students in 3 different English schools
- Conduct further case studies with 15 students in 1 English school
- Conduct evaluation activities outside the main evaluation protocol with 18 students in England,
- Screen all students at the start and end of the main evaluation period
- Set up accounts and for all students on the iLearnRW system and initialize their profiles
- Conduct individual interviews with over 55 English students at several key points in the evaluation in English schools
- Set up tablets for each student and teacher
- Provide further feedback to developers, test and release updates
- Develop and refine guidance for system usage

The timing divergences were due to the realities of working with schools as well as delays in delivery of the software caused by the late discovery of software bugs.

In addition, we were able to collect more observational and interview data from 55 English students at three points in the evaluation period:

- Following a mid-term 1-week break
- Preceding the period of independent study
- At the end of the evaluation period

The data collected will be analysed over the next three months and the results will be reported in deliverable 7.3.

2. Methodology and Evaluation Activities

2.1. Participants

2.1.1 Design-Based Research – UK

DBR evaluations involved both *specialist* and *non-specialist teachers* as well as *children who received additional educational support for literacy* within our target age group. This specifically involved teachers and pupils from three different UK state primary schools, the pupils from one of the schools (School B) had participated in the previous iteration (reported in D7.1) and the others were schools where two of the specialist teachers who participated in the expert appraisal (reported in D7.1) worked (School C and School D).

The schools varied in their profiles and teaching approaches.

- School B is located in an affluent area of south-west London, has low numbers of children at the early stages of learning English and just above average numbers of students with special educational needs. This school was judged to be outstanding by Ofsted.
- School C is located in a deprived and ethnically diverse area of east London, has a high proportion of children at the early stages of learning English and an average number of children with special educational needs. This school was judged to good by Ofsted.
- School D is located in a deprived and ethnically diverse area of north London, has high numbers of pupils with English as an additional language and above average number of student with special education needs. This school was judge to be good by Ofsted.

The schools chose slightly different approaches to supporting the children with dyslexia and other literacy difficulties due to the variations in how far behind their classmates these students were. School C followed a more differentiated approach to literacy learning where children who were severely behind in literacy within year 6 (top year) were typically removed from lessons to participate in a booster group if it is deemed that they would be unable to access the material. A similar approach was followed in School D where the children received regular individualised intervention sessions (i.e. multiple times a week). Whereas the teacher from School B described their inclusive approach to learning where all of their lessons and materials were designed to be ‘dyslexia-friendly’ to ensure the lesson as a whole was accessible to everyone, starting everyone at the same point and then the brightest students are set less-structured more challenging work to move onto. However, within this school lower-ability children did still receive some additional less frequent (i.e. once a week) individual interventions, but as the screening scores reflect the children receiving these interventions from School B were not as low ability as the children from the other two schools.

All of the girls from School B had been flagged up as not making sufficient progress in their literacy by the school and three of them received a writing intervention session with an outside specialist who visited the school once a week. They would also sometimes receive additional support within class from an LSA and were likely to form part of a smaller reading and writing booster group run later in the year by the class teacher during assembly time. The boys from School C were all receiving what is known as ‘Wave 3 intervention’, which is a targeted and individualised literacy intervention for pupils who are working at a level much lower than that expected of their age group. This involved the children coming out of class four times a week for 30 minutes to attend a 1:1 literacy intervention with an experienced specialist teacher. The children from School D were also receiving ‘Wave 3 Interventions’, M4 attended 4 intervention sessions a week for 40 minutes and F5 and M5 each attended 2 sessions a week for 30 minutes. The intervention sessions at this school were undertaken by

a fully qualified teacher who was currently retraining as a specialist dyslexia teacher. Table 1 provides an overview of the participants involved, their demographics and learning profiles.

School B	School C	School D
Class Teacher (Year 5) (non-specialist)	Specialist Dyslexia Teacher (experienced)	Specialist Dyslexia Teacher (trainee)
F1 - age 9, hearing difficulties, suspected dyslexia F2 - age 9, dyslexia diagnosis F3 - age 9, processing and memory concerns F4 - age 9, attention and memory concerns	M1 - age 10, gaps in phonics learning due to irregular school attendance M2 - age 10, suspected dyslexia, English as additional language M3 - age 10, social and communication difficulties	M4 - age 10, suspected dyslexia F5 - age 9, suspected dyslexia M5 - age 10, suspected dyslexia, gaps in phonics learning due to starting formal schooling late

Table 1 Overview of participants (F = female and M = male) including description of difficulties provided by their teacher

2.1.1 Programme Evaluation and Case Study Research – UK

The participants involved in the programme evaluation in English schools comprised of teachers and students at 4 schools. 60 children participated in the main programme evaluation activities. The mean age was 10 and 65% were male. 28 children who were not screened and used the system under observation only. 5 children dropped out during the course of the evaluation. The reasons included moving to a different school, parents changing their mind as in one case the child's own decision not to continue. Additionally, 5 teachers participated (2 specialist dyslexia teachers and 3 teachers provided by Dyslexia Action).

From this cohort, we further chose to focus more closely on a subset of children to support the case study research. We focused on one of the four schools involved and gained consent from 9 parents to involve their children in this further investigation. However, given school activities organised for 3 of these children over three out of four weeks of the case study observations, we ended up with a total of 6 cases.

2.1.1 Programme Evaluation – Greece

80 students were recruited in the in the final evaluation between 9 and 11 years. All children were attending special education classes in 10 primary schools in Ioannina, Greece and 67% were male. In addition, 15 teachers participated; 11 special educators and 4 teachers of the mainstream classes who were interested in observing the whole application, aiming to understand and be familiar with new technology applications in education, generally.

62 students composed a teacher-guided group and 18 a non-guided control group. They were all formally diagnosed with dyslexia and/or reading/writing difficulties and for all have been given consents by their parents to be involved in the evaluation phase.

2.2. Technology

Each tablet was set up manually as follows:

- Wi-Fi access

- Google account under Dyslexia Action domain shared across all tablets
- Several apps were necessary to install before iLearnRW apps could be deployed
 - Meraki to control the tablet remotely and deploy iLearnRW apps
 - AppLock (free version) to lock down access to specific apps such as Gmail
- Each tablet was connected to the free Meraki account for management
- Password was set up on App Lock
- Installation of non-Play-store apps was enabled
- iLearnRW apps were downloaded over Meraki and installed (alternatively, on limited internet connections, apps were copied from an SD card or a connected computer)

After the iLearnRW apps became available to distribute through the app store, they were gradually replaced on all tablets with the Play Store version. This enabled us to push updates to the software remotely, something which was not possible through Meraki.

It was possible to send messages to the tablets through Meraki but these turned out to be easy for the students to dismiss.

2.3. UK Implementation

The evaluation consisted of two parts. The **first** took place from September until December 2014 in three English schools and was concerned about the contextual uses of the software in specialist teaching, although it is noted that due to prototype instability this phase ended up forming part of the formative evaluation (see Section 2). The **second** was concerned about the effectiveness of the software in motivating learners and improving their reading skills.

2.3.1 Design-Based Research (DBR)

The first iteration of the DBR was reported in D7.1. During the **second iteration**, our goal was to develop supporting materials for the game. We began by observing existing literacy sessions at each of the schools to understand the pace and learning targets set by teachers. The sessions were observed by one researcher who sat at the back of the class and took extensive written notes throughout. In School B the researcher observed both a literacy class as well as a guided reading class. In Schools C and D the researcher observed an intervention session with each of the children.

A researcher also conducted interviews with each of the teachers to find out about their current teaching practices, use of technology and backgrounds of the participating children, and these interviews lasted approximately 30-40 minutes. The children were also interviewed by a researcher to find out about their previous experience of technology and attitudes towards literacy. All of the interviews were audio recorded after appropriate consent was obtained (from the teachers, children and their parents) and transcribed. Lastly initial try-out sessions (lasting between 15-20 minutes) were undertaken with both the teachers and the children to give them an introduction to the iLearnRW application, with the teachers from School C and School D each taking a tablet home with them to allow further exploration of the application. Again a researcher took written notes during each of these try-out sessions.

At the start of **Iteration 3** each of the children were given a tablet which had the iLearnRW game pre-installed on it and was set up for them to use. The two specialist teachers were also given their own tablets with the game on as well as the adapted supporting materials which included user guides generated from Iteration 2, mini-game recommendations and lesson plans whose importance was highlighted in Iteration 2.

The procedure for the field trial was agreed with each of the teachers in advance to fit in with the literacy teaching setup within each individual school. As the children in School B were not receiving the same form of intensive individualised intervention sessions as the other schools it was decided that the children would participate in weekly half hour sessions facilitated by a researcher where they would spend time playing the game in school and they would then be allowed to take the tablets home to use in between these sessions. Within School C and School D the children used the game during their regular literacy intervention sessions when it was deemed to be appropriate by the specialist teacher. The children in School C were also allowed to take their tablets into the classroom and were allowed to play on the game during some of their literacy lessons. All of the children's interaction with the games were recorded in the system logs.

The children were then able to use the tablets for a period of 8 weeks. Towards the end of the trial a researcher observed and audio recorded a session during which each of the children used the iLearnRW game (excluding 1 child who was absent). The length of time that the children used the game for during these observations ranged from 5-20 minutes. Lastly each of the teachers and children were interviewed about their experiences of taking part in the field trial. The teacher interviews ranged from 20-40 minutes and the children's interviews ranged from 10-15 minutes. Again all of the interviews were audio recorded and transcribed.

2.3.1 Programme Evaluation

a. Timetable and Procedure

Due to school schedules and timetable changes, the evaluation activities started over the course of three weeks in English schools from mid-January to early February. Therefore the dates below are relative to the total of 22 weeks of the evaluation. Also, due to constantly fluctuating time tables and differing holiday schedules at the schools, not all the schools were able to take part in all activities in the same relative times.

Initial screening and introduction to project and technology	Week 1
Start of group teaching sessions with students using technology	Week 2
Mid-term interviews with students	Week 5-6
Start of independent technology use at home	Week 11-12
Interviews with students	Week 11-12
Final screening and interviews with students	Week 20-22

The original plan was that after the initial introduction of the iLearnRW apps to the teacher, a technically adept member of the iLearnRW team would attend the first few sessions to make sure the software was functioning appropriately and any technical questions by the students would be answered. However, the apps required more updates than anticipated and experienced many other issues such as those with Wi-Fi connectivity and accounts that technical support was required in almost every session. As a result, most group sessions included two evaluators, a literacy teacher and a member of the iLearnRW team providing both technical and pedagogical support.

The English implementation included two steps: teacher/researcher led sessions at school followed by independent technology use at home. For the first 10-12 weeks (depending on the school), the key evaluation activity was a combination of weekly school sessions in small groups of 5 students combined with optional home use of the technology. The school sessions were scheduled to last 30 minutes but occasionally technical or other scheduling difficulties (e.g. preparation for tests and school plays) required us to extend them, or cut the sessions short. During the sessions the researcher/teacher fulfilled multiple purposes including:

- Support students as they used the technology with additional language input (including explanations of metalinguistic terms and help with difficult words)
- Understand barriers involved in students' use of the technology to provide further feedback to developers
- Scaffold students' use to the technology in a gradual way and expose them to new features
- Provide technical support to the students in the use of the technology as well as android tablets
- Provide instruction in certain metacognitive and metalinguistic aspects to fill gaps in the delivery of school curriculum. For instance, not all children had been taught what a prefix is.

For the last 10 weeks, all students used the tablets independently without weekly support and encouragement from the evaluation team. Students were given written instructions (see Appendices for list) as well as contact details to receive additional technical support.

b. Screening Measure

All children participating in the formal programme evaluation were administered a short screening. This consisted of a short dictation administered in small groups, followed by a 1-page assignment with identifying parts of words on the page. This was followed by a brief individual interview where the student read out words with the interviewer noting their correctness. Several children who missed the group session were the entire screening individually.

The portion of the screening focusing on the identification of syllables, suffixes and prefixes was introduced with a brief checking of understanding of the concept and follow up explanation to ensure that the children focused on the task. This was necessary since the curriculum proceeds at different paces in different schools and also because our age range included students across 3 different school years. The majority of children had been exposed to the concepts but were not able to immediately respond to the question of 'What is a suffix?' (prefix, or syllable).

Due to drop out we have administered the final screening to fewer students than the initial screening.

The results of the initial screening were entered into the iLearnRW system to help initialise the phonic profile essential for the most efficient functioning of the system. The final screening was only administered in the final days before the deliverable (see above) and therefore will be entered into the system at the start of the evaluation data analysis phase.

Samples of the screening questionnaire format are included in deliverable 7.1 and a complete list of the words and tasks used is provided in this deliverable as an appendix. The process of word selection for the screenings was also described in deliverable 7.1.

c. Usage logs

The main source of data for the evaluation of iLearnRW as described in deliverable 7.1 were the comprehensive logs collected by the game and the reader application. Every time the student launched the game, logs were collected for every action they performed in the game. In particular, the following was logged:

- Start of session
- Settings and personal preferences
- Suggested activities
- Launch of activity
- Words presented
- Success or failure of word interaction

Sample of raw log data is provided in the appendices.

The log data was processed and made available to the teacher and the evaluation coordinators to enable the monitoring of evaluation activities.

Sample of pre-processed and visualised data are provided in the appendices.

During the evaluation period, the logs have recorded several thousand hours of game play across more than two hundred students. The detailed analysis of the logs will be conducted during the final phase of the evaluation and reported in deliverable 7.3.

Note: We expect some level of unreliability in the log data due to interruptions in the logging process during internet connectivity. This will mainly affect log data regarding the ends of sessions but also in some cases missing logs of activity play due to connectivity errors in the log transmission.

d. Student Interviews

The evaluation plan included interviews regarding game usability and evaluation of gameplay. However, during the process of the evaluation in English schools and after a preliminary survey of the logs, it was decided to expand the questionnaires to include information about students' perceived learning gains and self-reported usage patterns.

The purpose of the student interview was further expanded to include suggesting further activities to the student and reinforcing the purpose of the evaluation. The middle interview was also used to reiterate the assignment.

In total, three questionnaires were administered to each student as part of a structured interview. One as part of a group interview and two as individual one-on-one interviews with one of the evaluators. The timing of the interviews was

- Interview 1: Following a mid-term 1-week break (small groups)
- Interview 2: Preceding the period of independent study (individual)
- Interview 3: At the end of the evaluation period (individual)

In one school, the interviews were combined with design-based research activities (see below).

The protocol follow in these interviews was as follows:

Small group interview

Interviewer works with a group of 3-6 students and asks them questions following a structured questionnaire.

Interviewer makes notes about the answers and when appropriate takes a poll of how many students of the group engaged in the activity.

Interviewer makes notes about the group responses, noting student numbers as indicated on the protocol.

Individual interviews

Interviewer works with an individual child following the structured interview protocol. The child answers are noted down by the interviewer during the interviews.

A small portion of the interviews was recorded as part of the design-based research described below. Special permission had to be sought for that from parents.

Not all questions were appropriate to every single child depending on their concrete circumstances. Sometimes the answer was already known to the interviewer (e.g. when asking about technology at home) and in that case was merely noted. Also, frequently, the child would answer a question as part of an answer to an earlier question.

Full text of the structured interview questionnaires is included in the appendices.

2.3.1 Case Study Research

The programme evaluation followed a group intervention model where children are placed in small groups to receive support. Within this context, we conducted a focused case study whose aim was to understand the socio-emotive dynamics that reinforced children's motivation or became barriers to their learning experience. To answer this question, it was necessary to document and observe children's interactions amongst each other at a micro level while using the game. We achieved this by video and audio recording each session for a period of four sessions (over four weeks) yielding a total of 2.5 hours of recorded data. Additionally we took observation notes during each session of possible incidents for later probing. Unfortunately due to technical difficulties in running screen capturing software on Android tablets, we were unable to record children's concurrent game play. Instead, we rely on the logs to provide us with the game play context underpinning children's group interactions. We note that our unit of analysis is at the group level.

To avoid promoting socially desirable responses and behaviours, it was important to gain children's trust and to become accepted as part of the social context. Therefore, one of the researchers involved in the case study attended four sessions before the data collection began, whereas the second researcher involved attended sessions on a biweekly basis. After the data collection ended, researchers viewed the videos and their notes with the goal to identify critical incidents regarding collaboration, conflict and peer interaction. These critical incidents were probed during post-interviews with children (averaging 30 minutes each) alongside questions about the specific games that allowed us to understand how social interaction was facilitated by shared game preferences amongst the children. Video recordings and interviews were transcribed for subsequent analysis triangulating video observations with interview responses from the children.

2.4. Greek Implementation of Programme Evaluation

The Greek format of the programme evaluation followed the same principles as those described in section 2.3.1. The key differences followed from the different school environments in which the evaluation was conducted.

- The Greek student groups were divided between those who were receiving in-school intervention on multiple days a week with a specialist teacher who agreed to use the software as part of 3 sessions a week for 30-45 minutes and those who were using the tablet completely independently.
- During the Greek sessions, only 1-3 students took part at once depending of the school arrangement.
- Since most students were observed by teachers during their use of the tablet, a questionnaire was administered to teachers at the end of the evaluation rather than students

2.5. Summary of Deviations from Original Evaluation Plans

There were several factors that caused several changes in the implementation of the activities. These changes were:

- Changes in timing of parts of the evaluation
- Expansion of the support provided to students and teachers
- Expansion of qualitative evaluation data collected

The factors that led to these changes were:

- Last minute changes in school timing and availability
- Continuing updates to software functionality
- Ongoing monitoring of expansion process

See details about all these factors in the conclusions and recommendations.

We were also able to use the system under various observational conditions with a much larger number of students than originally intended. This was due to a lot of interest in using the system among teachers. While not all these students were fully screened and debriefed, we were able to incorporate observations and lessons learned from them into the evaluation.

3. Conclusions and Lessons learned

The analysis of the data collected is described in detail in deliverable 7.3. This section described lessons learned during the process of collecting the data. While we were successful in collecting the data required for the evaluation of the projects, we have learned a number of lessons that will be relevant to future.

3.1. Working with students

The key challenges of working with students were a result of two factors:

- Classroom management
- Home vs. school work interaction

The issue with classroom management derived from increased support needs due to the continuing developments in the software. While groups of 5 appear well-suited to the format of working with students who need some support and can work independently, this becomes much more difficult when issues become more frequent and support needs escalate. The issue was not just technical difficulties but also the lack of full implementation of certain features of the apps meant to provide guidance and scaffolding to the students using them. This meant that students did not become confident users of the apps until many weeks into the project.

This was exacerbated by the fact that we were working with children with possible attention difficulties as well as many disaffected children with mild behavioural problem. This came to prominence when our session were scheduled instead of Physical Education. This occasionally led to severe behavioural problems with one boy accusing us of making him miss the only thing he liked in school.

We recommend smaller groups or more support staff available when working with prototype software. We also suggest that help and in-software guidance are made a priority during development.

This also had a knock on effect on the timing of teaching sessions. While 30-minute sessions seemed ideal when everything was working well, they occasionally seemed too short when delayed and interrupted by technical difficulties.

The other issue disrupting in-school evaluation activities was student forgetting to bring tablets from home or bringing their tablets completely discharged. Sometimes only 1 student out of 5 had their own functioning tablet available.

We were able to make up for the lack of tablet with bringing spare tablets to every session. However, this became more difficult over time as these tablets had to be used to replace broken tablets. In some case, students had to work in pairs.

This situation had an even bigger impact on keeping the iLearnRW apps uptodate on all tablets. Particularly before they were distributed through the Play Store. Some children had out of date apps on their system for several weeks.

We recommend that projects requiring students to take tablets home include additional tablets for school work to avoid the issues with forgotten tablets. This would also necessitate distributing the

apps via the Play Store for ease of updates (or not updating the apps during the evaluation process) and setting up each tablet with an individual Google account for ease of monitoring.

3.2. Working with schools

Working with schools presented its own set of challenges. The biggest issues were with room allocation and scheduling changes.

All schools were extremely helpful and accommodating but the evaluation was not a priority for them. Students in Year 6 were preparing for SAT exams and also took part in end-of-year activities such as plays and school trips. While we presented each school with a schedule of evaluation activities, they often only let us know about schedule changes at the last minute or not at all. In three schools, evaluation activities had to be moved to another week due to a school trip about which we were informed only the week before. In one school, we arrived only to find two thirds of the students about to depart for a field trip to a theatrical performance. In another school, the teacher who kept the tablets locked in preparation for giving them to students to take home, went away with another class and we had to postpone giving the students their own tablets by a week. In one extreme case of miscommunication, we arrived at a locked school only to find out that their holidays are different from others – this despite confirming with the deputy-head the date of arrival.

There is no concrete recommendation made based on these experiences other than to expect similar issues. We expected many of these and tried to mitigate for them by frequent explicit communication with the schools, yet we were unable to prevent the issues.

We also experienced frequent lack of space and changes of location in schools. This was also expected and unavoidable due to the demands on space in all schools. However, it still presented a disruption.

3.3. Hardware

While the hardware selected was sufficient to run all the iLearnRW software and other software necessary for the evaluation, we experienced a number of hardware malfunctions ranging from deteriorating screens to faulty charging.

Many tablets also developed a fault where all sound disappeared until the tablet was restarted. However, this solution was not obvious and therefore some children missed several weeks of play.

We also experienced slightly more breakages of tablets by students than expected from reports on previous evaluations. Most frequent of these were broken screens on 5 tablets. One student spilled water on their tablet and one student lost their tablet. Another student left their tablet while visiting relatives abroad.

We budgeted for 10-15% tablet loss and breakages and as a result we expected able to replace all broken tablets. This was partly facilitated by the dropping out of several students and their tablets becoming available.

However, we did not plan or budget for lost chargers and charging cables. We were able to provide spares from broken tablets and only needed to purchase several cables. However, we recommend space be made in the budget for that.

We also had to purchase headsets for work in groups to prevent noises from the game disturbing other students. Since this was not anticipated in the budget, we purchased the cheapest possible headsets resulting in many breakages.

The setting up of tablets took a considerable amount of time and we recommend that more time and human resources are budgeted for this activity.

We recommend that any future project includes flip covers protecting the screen in the budget for all tablets. This could increase the total budget of the hardware by as much as 10%. We purchase cheap neoprene sleeve cases for tablets used by teachers but they only protect the tablet during transport and not during use. Most screen breakages were reported to have happened accidentally while the tablet was unlikely to have been put in a sleeve.

3.4. Software development and deployment

Android proved to be an ideal system for this project in many ways but it was limited in the how much remote control we had over each tablet. While we controlled each tablet with Meraki, since we used the same Google account on all tablets, we were unable to identify which tablet belonged to each individual child. If more time was budgeted for tablet set up we would recommend setting up an individual Google account for each child. This would also enable them to personalise their tablet without potentially compromising other children's tablets. However, this will also lead to issues if the project uses apps sold through the Google Play store. Using the same account on all tablets only requires one purchase whereas separate accounts would require an individual purchase for each tablets. This would represent a small financial burden but a much bigger administrative burden since credit card details would have to be entered and then deleted for each individual account.

Due to technical difficulties, the iLearnRW software was not available through the Google Play store until after the start of the main evaluation activities. We used Meraki or manual installation to deploy the apps initially. This had the advantage of not always having to rely on internet connectivity but this was outweighed by the disadvantage of not being able to provide updates to the apps when they became available to all tablets at once or even to tablets remotely and automatically.

This was made further difficult by the size of the game app which at one point reached over 300MB. This meant that download and installation over the internet could take a considerable amount of time over slower school connections. We frequently had to supplement school connectivity with 3G.

School connectivity (even with back up 3G connectivity) also proved to be an issue when using the games. While we were able to connect all tablets to the school Wi-Fi₃₃, this was not always reliable. In some schools, we had to move rooms with an impact on connectivity. Frequently students could not start playing because they could not log in and logs could not be sent or suggestions for further play retrieved. This resulted in many errors during game play and possibly makes the log data incomplete (see above).

We recommend that all software requiring logging in is asynchronous, where logs are collected offline and synched with the server when connectivity is available. While this introduces other possible issues, they would be less disruptive to teaching and usage.

Variable Wi-Fi₃₃ connectivity also made updating the iLearnRW game slow and very unreliable over the Play Store or Meraki. Particularly, trying to update up to 5 tablets at once put a strain on the

Date: 2015/06/30

Project: ILearnRW

Doc.Identifier: D7.2 Evaluation Testbed Implementation_FINAL.docx



Wi-Fi. At one school, several times, the evaluation team had to arrive several hours earlier, collect all tablets from students, take them to a location with internet connectivity and update the game.

We recommend that app size is given more attention in prototype development and options are explored for allowing update without redownload of the entire software package.

4. Appendices

4.1. Screening Test complete data

4.1.1 Initial screening English

Read out these words

fit, nice, act, lots, hand, car, little, unit, upon, tiny, very, spin, strike, stress, rim, kitten, plant, melted, branch, class, wishes, baby, soft, wind, income, create, path, milk, seven, wash, shrug, keeping, tax, jazz, unit, planned, cutting, force, career, branch, guilty, decide, maybe, decided, pricing, sentence, walked, timed, darkest, wood, tooth, bridge, construction, nearly, central, provide, possible, tackle, county, fight, stories, church, firm, brief, traditional, author, magazine, famous, engine, type, receive, player, castle, knife, science

Split word into syllables

little, baby, carry, student, poem, weekend, capture, family, ability, recommendation, animal, benefit, majority

Dictation

heat, frown, friend, magic, punish, exchange, important, away, paper, combined, shoot, produce, stable, church, decision

Split the suffix

wanted, classes, homeless, useful, helping, fitting, dropped, sadly, possibility, loved, typing, harder, calmest, collection, comfortable, global, happiness, heavier, submitted, artist, special, mutual, politician, editor,

Split the prefix

unknown, inside, appear, success, difficulty, effective, understand

4.1.1 Final screening English

Read out these words

bike, sit, tact, pots, land, cell, lucky, shiny, until, music, better, spill, spike, fresh, trick, letter, spelt, slide, glass, bench, washes, away, gift, kind, include, react, maths, think, never, watch, brush, booking, six, puzzle, parent, hissed, kidding, storm, cheer, bunch, dirty, delay, pride, forced, rising, silence, worked, named, fastest, smooth, hook, judge, action, beard, formal, comfortable, terrible, circle, ground, might, cities, purchase, third, brief, search, conclusion, imagine, machine, nervous, ceiling, style, amount, listener, knock, scene

Split word into syllables

better, paper, attack, human, flying, lemon, picture, enemy, information, unfortunately, imagine, military, occasionally

Dictation

cream, brown, frame, logic, publish, example, divorce, player, power, planned, tooth, notice, double, nurse, attention

Split the suffix

asked, bushes, endless, careful, looking, getting, stopped, blindly, credibility, faced, hoping, darker, highest, reflection, remarkable, formal, nastiness, prettier, happened, tourist, social, virtual, musician, investor

Split the prefix

unlikely, inform, commander, advocate, difference, efficient, undertake

4.1.1 Screening test Greek

Διάβασε φωναχτά τις λέξεις που βλέπεις παρακάτω:

κούκλα, κήπος, μαγνήτης, μήνυμα, συμμαχία, δρομος, αφίσα, λαιμός, σώμα, λάμπα, δέμα, θέμα, κτήμα, πτηνό, δρόμος, θρύλος, όχθη, λαϊκός, μαϊμού, καημός, σαΐτα, ξένος, υψώνω, στρώμα, σκράπα, σπρώχνω, φρύδι, χρόνος, σπιτάκι, γραφίδα, στρατηγός, στρίψιμο, σφραγίζω, ξεσπώ, κάστρο, αφρός, εχθρικός, χωρίστρα, κρεμάστρα, αστραπή, εύκολος, λεύκωμα, δεύτερος, μαϊντανός, προϋπαντώ, χορηγός, βήχω, χάνω, δένω, κάνω

Χώρισε τις λέξεις σε συλλαβές, τραβώντας κάθετες γραμμές όπως στο παράδειγμα: π.χ. κορίτσι -> κο | ρί | τσι

ξανά, ξηρά, έξω, ουρά, χάος, νέος, ζωές, κούκλα, γλύπτης, σταθμός, άστρο, εχθρός, φιλία, ηρεμία, φιλία, αγκαλιά, ταΐζω, είδος, λείπω, φορείο, γάιδαρος, λαϊκός, μαϊμού, κραυγή, χορεύω, παλεύω, άλφα, καρφί

Κύκλωσε το πρώτο μέρος της λέξης όπως στα παραδείγματα: π.χ. αντίσταση -> (αντί)σταση, αντιπαθώ -> (αντι)παθώ

αντίθετος, υπερβολή, υπογραφή, αναζητώ, ψιλοβαριέμαι, υποκινώ, πρωτοβρόχια, υπερβάλλω

Κύκλωσε την κατάληξη της λέξης, όπως στα παραδείγματα: π.χ. παιδάκι -> παιδ(άκι), ποδόρα -> ποδ(άρα)

κοπελίτσα, σακουλάκι, κοριτσάρα, ατακτούλης, τρέξιμο, εξυπνάδα, αρρωστιάρης, μοσχαρίσιος, δυναμώνω, χορεύω, τεχνίτης, πορτιέρης, ταξιτζής

Κύκλωσε την κατάληξη της λέξης, όπως στα παραδείγματα: π.χ. άνθρωποι -> άνθρωπ(οι), λύση -> λύσ(η), τρέχεις -> τρέχ(εις)

πωλητής, πόλη, γκρινιάρηδες, μεγάλου, αγάπησα, γράφουμε, αγαπιέμαι, χαιρόταν, παρών, καναπέδες, παπάδων

4.2. Interview questionnaires

4.2.1 Half Term Student Questionnaire

e. Playing the game

Who played the game over the break and how many times?

no 1x 2x 3x more than 4x

Which games did you like playing the most?

How many photos did you get?

What do you remember about the story?

Which ghosts did you make friends with? What did you learn?

What did you like about the games?

What would you like to see improved?

f. Using the reader

Who used the reader over the break?

no 1x 2x 3x more than 4x

What did you read?

Ask whole groups. Write down number of students who raised hand in answer to questions. Make notes on any other answers.

Group No: _____

4.2.1 Pre-Independent Study Interview

1. Tell me how you've been using the game?

Supplementary prompt questions: When did you play it? How long have you played it? If you have not played it, why? What did your parents think? What did your friends think? What can we do to make the play experience better? Did you enjoy getting the rewards? Have you used the reader?

2. Has it improved your reading?

Supplementary prompt questions: Did you know what you're trying to achieve? How did you deal with words you didn't know? How did you deal with things you didn't understand?

3. What are things you need help with in the game?

Supplementary questions: Did you know you can now: 1. do things more quickly (list things)? 2. Mute music in Music Hall? 3. Skip trains in Dispatcher. Did you know you can get back to history?

4. How will you use the game from now until June?

Supplementary questions: Will you take out the tablet at least 2x a week? When will you play it (school/home)? What will you do when you don't have Wi-Fi (take it to school)? What will you do to get more photos?

4.2.1 Final interview

Tell me how you've been using the game?

Where did you use it?

How did you use the games?

How/when did you use the reader?

2. Has it improved your reading or writing?

Do you feel more confident in your reading now?

How did you feel before?

How did using this change your writing? Speed? Spelling? Composition?

Has it changed how you read?

Has it changed what you do in school?

Any other comments about changes in reading / writing?

Did you learn anything new about using tablets for learning/study/school?

Do you think you will look for more learning games?

Phone: iPhone Android (note brand) Windows phone
Own Shared with sibling Family other

Tablet:	iPad Mini	iPad (big 10inch)	Android 7"	8"	9"	10"	other
	Own	Shared with sibling			Family		other

What do you think will change?

Which of the games did you learn the most from? And what did you learn?

Any other comments/suggestions?

- Activity Reference Guide.docx
- Character Reference Guide.docx
- Character Short Descriptions.docx
- Characters Language Matching.docx
- Game Quick Start Guide.docx
- Game User Guide.docx
- How to Install own apps on tablet.docx
- How to Update Words Matter.docx
- How to use MoonPlus Reader.docx
- Language Explanations.docx
- Language Glossary.docx
- Reader User Guide.docx

4.4. Mid-evaluation communication with parents

Where we are

We are now about to conclude the classroom sessions of the evaluation.

All **children have tried** with supervision:

- Playing the game and reading using the reader
- Checking their progress and seeing rewards
- Using the suggested next game feature

What will happen next

Until the end of June, all children will be **using their tablets independently** at home. When they return their tablets, every child will be given **a short screening test** and answer a few questions.

We may also **contact some parents** who gave us their contact details.

Home study assignment

During the independent study, children are expected to:

- **Play the Words Matter game** at least twice a week for **15-20 minutes**.
- **Get photos for all characters** in the game. Each photo means they have practiced the given skill.
- Read at least two short stories a week using the **Words Matter reader**.

All gameplay and reading are being tracked by the project. At the end of the project, we will give **small prizes** to children who 1) play the game the most, 2) make the most progress, 3) use the reader the most.

Key dates

- Final classroom session **Fri, 1 May**
- Independent study May, June
- **Tablet return** and final screening **Fri, 26 June**

Key contact details

If you have any questions about the evaluation or the tablet, please contact.

name, email

phone number (please leave a message)

Key tablet details

Play store account

Sometimes, this is required to get updates for the game.

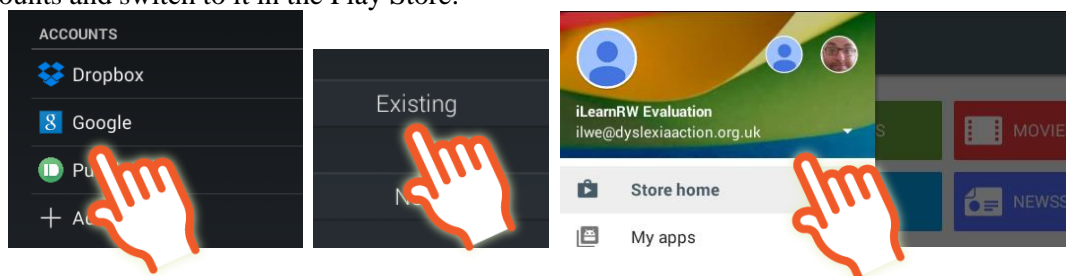
Account name: email

Password: password

Warning! This account is shared with all others in the evaluation. **DO NOT** use this account to purchase any apps or games. Others may see your payment details and you will not be able to download the games anywhere else.

To install your own apps

Before you install own apps, you should add your **own Google account** to the tablet via Settings > Accounts and switch to it in the Play Store.



To install Words Matter apps on your own tablet

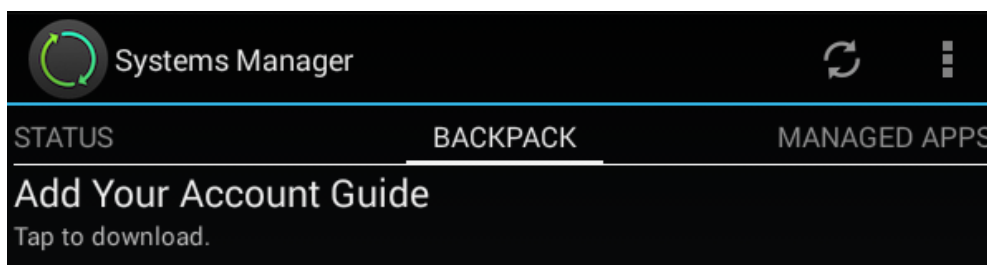
1. Add ilwe@dyslexiaaction.org.uk Google account to your tablet (as above)
2. Switch to that account in the Play Store and search for 'ilearnrw'
3. Install Words Matter and Words Matter Reader
4. Login details for game/reader: **Username:** ruchildname **Password:** cats

Note 1: Make sure to switch back to your own account before installing other apps or setting up a credit card.

Note 2: After adding this account, you may get 'Account Action Required' notifications. It is save to ignore these

Game manual and tips

Under Backpack in **Systems Manager**, you can find documents about the tablet and the game.



4.5. Sample of Raw Log Data

2015-05-04 14:51:05.0	nostudent	LOGGING_SCREEN	LOGIN		0	0	Connected
2015-04-30 19:47:08.0	nostudent	GAME_WORLD	SAVEFILE		0	0	
2015-04-30 19:46:52.0	nostudent	SERENADE_HERO	WORD_FAILED	nerve	2	51	
2015-04-30 19:46:52.0	nostudent	SERENADE_HERO	WORD_DISPLAYED	wander	2	55	
2015-04-30 19:46:39.0	nostudent	SERENADE_HERO	WORD_SUCCESS	smell	2	5	
2015-04-30 19:46:39.0	nostudent	SERENADE_HERO	WORD_DISPLAYED	nerve	2	51	
2015-04-30 19:46:31.0	nostudent	SERENADE_HERO	WORD_DISPLAYED	smell	2	5	
2015-04-30 19:46:30.0	nostudent	SERENADE_HERO	WORD_SUCCESS	warrior	2	55	
2015-04-30 19:46:23.0	nostudent	SERENADE_HERO	WORD_SUCCESS	smart	2	5	
2015-04-30 19:46:23.0	nostudent	SERENADE_HERO	WORD_DISPLAYED	warrior	2	55	
2015-04-30 19:46:15.0	nostudent	SERENADE_HERO	WORD_FAILED	dwindle	2	54	
2015-04-30 19:46:15.0	nostudent	SERENADE_HERO	WORD_DISPLAYED	smart	2	5	
2015-04-30	nostudent	SERENADE_HERO	WORD_SUCCESS	wash	2	55	

Date: 2015/06/30

Project: ILearnRW

Doc.Identifier: D7.2 Evaluation Testbed Implementation_FINAL.docx



19:45:58.0							
2015-04-30 19:45:58.0	nostudent	SERENADE_HERO	WORD_DISPLAYED	dwindle	2	54	
2015-04-30 19:45:51.0	nostudent	SERENADE_HERO	WORD_SUCCESS	smooth	2	5	
2015-04-30 19:45:51.0	nostudent	SERENADE_HERO	WORD_DISPLAYED	wash	2	55	
2015-04-30 19:45:42.0	nostudent	SERENADE_HERO	WORD_SUCCESS	swallow	2	57	
2015-04-30 19:45:42.0	nostudent	SERENADE_HERO	WORD_DISPLAYED	smooth	2	5	
2015-04-30 19:45:26.0	nostudent	SERENADE_HERO	WORD_DISPLAYED	swallow	2	57	
2015-04-30 19:45:24.0	nostudent	SERENADE_HERO	APP_ROUND_SESSION_START		2	57	Quest
2015-04-30 19:42:09.0	nostudent	GAME_WORLD	SAVEFILE		0	0	
2015-04-30 19:42:09.0	nostudent	MAIL_SORTER	ACTIVITY_PROPOSED		4	38	
2015-04-30 19:42:09.0	nostudent	MAIL_SORTER	ACTIVITY_PROPOSED		4	44	
2015-04-30 19:42:09.0	nostudent	MAIL_SORTER	ACTIVITY_PROPOSED		4	45	
2015-04-30 19:42:09.0	nostudent	MAIL_SORTER	ACTIVITY_PROPOSED		4	42	
2015-04-30 19:40:43.0	nostudent	MAIL_SORTER	ACTIVITY_PROPOSED		4	42	

Date: 2015/06/30

Project: ILearnRW

Doc.Identifier: D7.2 Evaluation Testbed Implementation_FINAL.docx

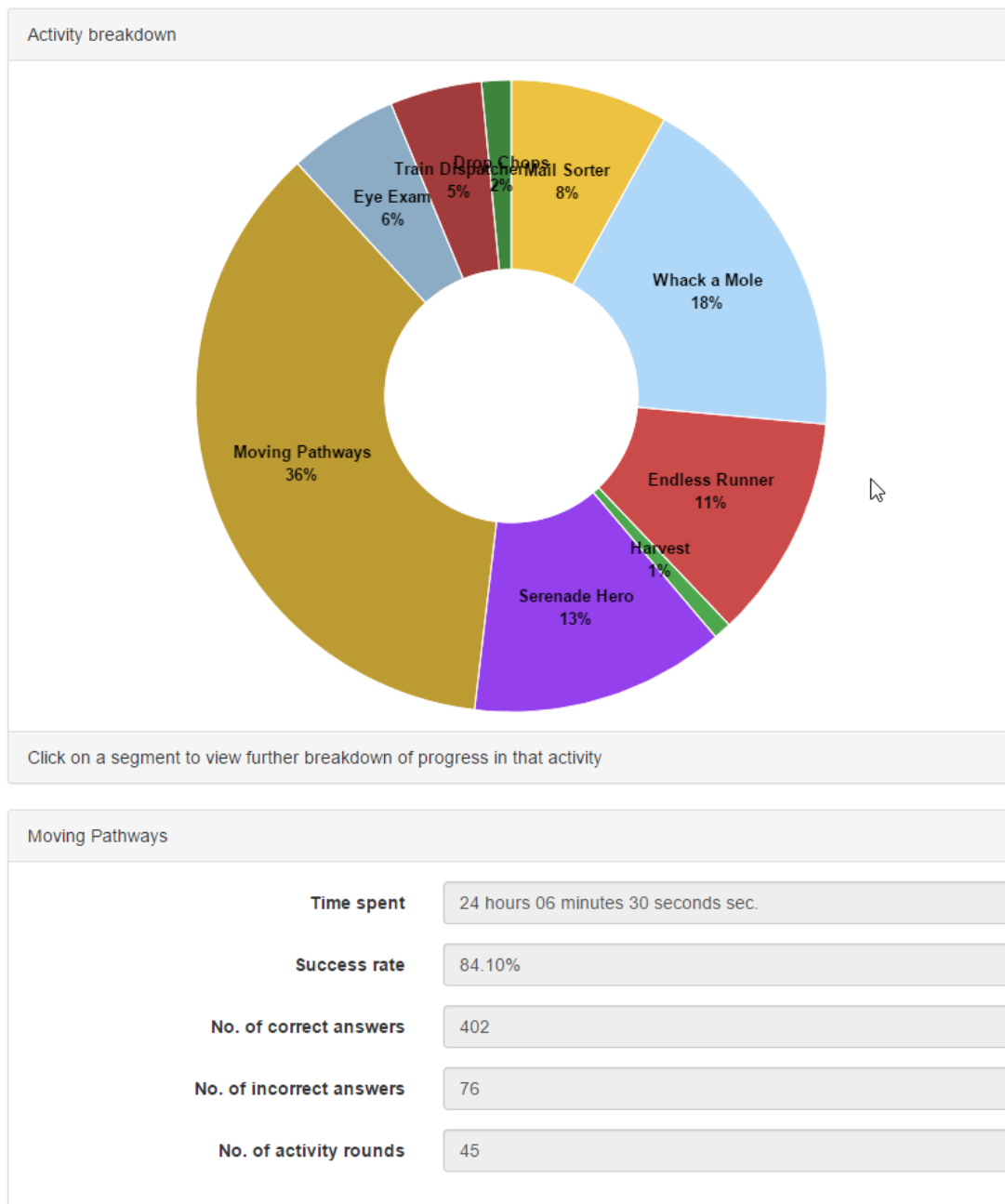


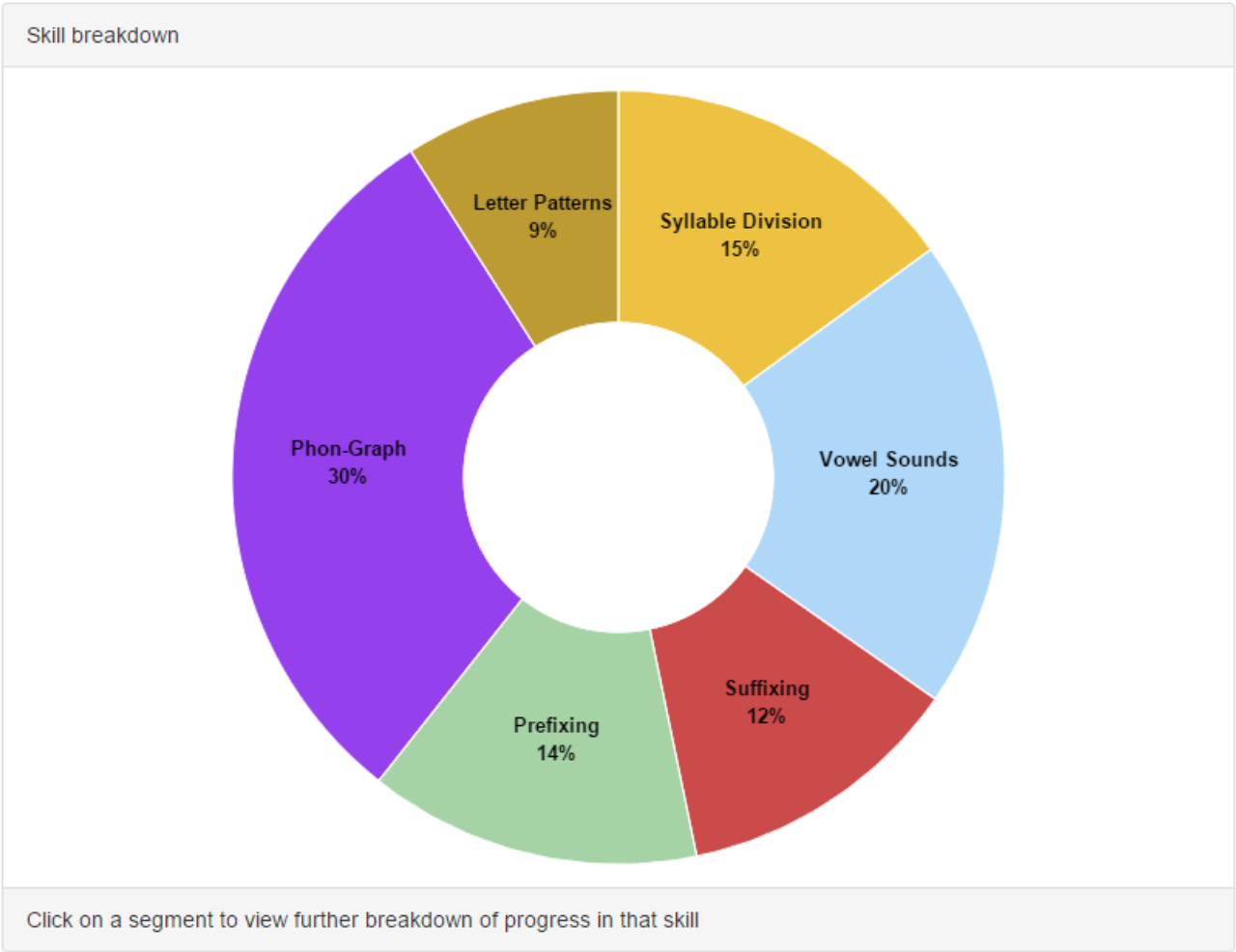
2015-04-30 19:40:00.0	nostudent	HARVEST	APP_ROUND_SESSION_END		4	43	
2015-04-30 19:39:19.0	nostudent	HARVEST	WORD_DISPLAYED	ridiculous	4	67	
2015-04-30 19:39:19.0	nostudent	HARVEST	WORD_DISPLAYED	drugged	4	29	
2015-04-30 19:39:19.0	nostudent	HARVEST	WORD_DISPLAYED	backed	4	30	
2015-03-25 14:50:07.0	nostudent	EYE_EXAM	APP_ROUND_SESSION_END		4	14	
2015-03-25 14:49:56.0	nostudent	EYE_EXAM	WORD_SUCCESS	hotly	4	16	ly
2015-03-25 14:49:30.0	nostudent	EYE_EXAM	WORD_DISPLAYED	hotly	4	16	
2015-03-25 14:49:26.0	nostudent	EYE_EXAM	WORD_SUCCESS	sickly	4	16	ly
2015-03-25 14:48:56.0	nostudent	EYE_EXAM	WORD_DISPLAYED	sickly	4	16	
2015-03-25 14:48:52.0	nostudent	EYE_EXAM	WORD_SUCCESS	treaty	4	14	y
2015-03-25 14:48:25.0	nostudent	EYE_EXAM	WORD_DISPLAYED	treaty	4	14	
2015-03-25 14:48:22.0	nostudent	EYE_EXAM	WORD_SUCCESS	manly	4	16	ly

4.6. Sample of pre-processed log data

Student name	Total time spent playing	Days played	Activities played	Skills practiced	Words seen	Success rate
nostudent	50 hours 59 minutes 59 seconds	2015-01-30, 2015-02-16, 2015-03-20, 2015-04-22, 2015-06-02	Logging Screen, Mail Sorter, Whack a Mole, Train Dispatcher, Serenade Hero, Endless Runner	0 EN, 2 EN, 4 EN, 1 EN, 3 EN	spelt, job, bring, -est, darker, finest, gamer, merest, abler, palest, purest, checker, payer, truest, easy, homing, moving, making, giving, casing, gaming, liking, crazy, wet, nod, six, fat, box, ten, fan, win, kid, dog, documentary, -es, does, byes, goes, owned, endless, evenness, grip, spite, staff, split, spy, chip, stress, snow, ship, stuff, snake, trip, stick, sneak, snap, spare, slip, stretch, space, thin, at, man, up, new, it, now, to, break, blow, black, block, brake, brush, brown, bride, brand, brick, brave, require, quarter, quiet, question, frequent, sequence, become, never, woman, giant, poem, via, liquid, equip, react, about	70.97% (88 out of 124)

4.7. Sample of visualisation of log data





Prefixing	
Time spent	56 hours 15 minutes 50 seconds sec.
Success rate	82.67%
No. of correct answers	458
No. of incorrect answers	96
No. of activity rounds	105

4.8. Final certificate of appreciation

