FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Gamification and Behaviour Recognition in Literacy Assessments for Sub-Saharan African Communities

Marcos Oliveira

Mestrado Integrado em Engenharia Informática e Computação

Supervisor: José Luís Moura Borges, PhD

Second Supervisor: Joana Polónia Coelho de Azeredo Lobo, Msc

July 23, 2015

Gamification and Behaviour Recognition in Literacy Assessments for Sub-Saharan African Communities

Marcos Oliveira

Mestrado Integrado em Engenharia Informática e Computação

Approved in oral examination by the committee:

Chair: Rui Carlos Camacho de Sousa Ferreira da Silva, PhD
External Examiner: Pedro Manuel Henriques da Cunha Abreu, PhD
Supervisor: Doctor Name of the Supervisor

July 23, 2015

Abstract

Sub-Saharan African countries present the lowest adult and youth literacy rates in the world. Simultaneously, they have led the growth of mobile communications worldwide, mainly done by smartphones. In this sense, mobile devices can play a crucial role in monitoring the literacy of this African region.

The purpose of this dissertation is to develop a mobile application to assess and monitor the degree of literacy of children and adults in sub-Saharan communities. In this context, the evaluation should be realized without the need to be performed by an individual specialized in the area, and using standardized and at the same time captivating tests. In order to keep the user committed and engaged to the process, it is intended to use gamification techniques. During the literacy assessment process, specific behaviours of the user interacting with the application may occur. Through these same behaviours, the application may identify and analyse the user profile, allowing the interface and usability aspects to be adjusted based on these aspects. Thus, this tool stands out from traditional literacy tests based on pen-and-paper.

The solution consists, initially, in conducting a literature review that covers the different tools and literacy assessment tests, the human-computer interaction for African users with low literacy rates, as well as gamification techniques and adaptive interfaces on mobile devices. The application will be developed according to an iterative process of user-centred design. Thus, after defining an initial strategy, iterations are expected through the processes of analysis and requirements specification, conceptual design and evaluation, until a final result is obtained. Due to geographical constraints, three first sessions of tests were conducted in Portugal, participating users aged between seven and twelve years old, having distinct levels of knowledge of the English language and different types of contact with mobile devices. The results show the efficacy of gamifying literacy in order to provide a captivating experience to the users and promising results about the adaptive interface.

It is expected that this tool will help improve the literacy rate in sub-Saharan Africa, enabling teachers and volunteers to diagnose and easily monitor the degree of literacy in these regions. It is also expected that it will stimulate the spirit of mutual help among community members, therefore, combating low level literacy in their home areas.

Keywords: literacy assessment systems, gamification of literacy, user behaviour recognition, user-centred design, adaptive interfaces, sub-Saharan Africa communities

Resumo

Os países da África Subsariana apresentam taxas de literacia juvenil e adulta mais baixas do que qualquer outra região do mundo. Simultaneamente, estes têm liderado o crescimento de comunicações móveis a nível mundial, sobretudo realizadas por *smartphones*. Nesse sentido, os dispositivos móveis podem assumir um papel fundamental no controlo da literacia desta região africana.

O propósito desta dissertação é desenvolver uma aplicação móvel para avaliar e monitorizar o grau de literacia de crianças e adultos das comunidades subsarianas. Neste âmbito, a avaliação deve ser realizada sem a necessidade de ser executada por alguém especializado na área, através de testes uniformizados e, ao mesmo tempo, cativantes. De forma a conseguir envolver o utilizador, pretende-se recorrer a técnicas de *gamification*. Durante o processo de avaliação de literacia, comportamentos específicos de interação com a aplicação poderão ocorrer por parte do utilizador. Através desses mesmos comportamentos, a aplicação pode identificar e analisar o perfil do utilizador, permitindo que a interface e usabilidade se ajustem conforme esses mesmos aspetos. Assim, esta ferramenta destaca-se dos tradicionais testes de literacia baseados em papel e caneta.

A solução passa por, inicialmente, realizar uma revisão bibliográfica que cubra essencialmente as diferentes ferramentas e testes de diagnóstico de literacia existentes, que estude a interação pessoa-computador para utilizadores africanos com baixo nível de literacia, assim como técnicas de *gamification* e interfaces adaptativas em dispositivos móveis. A aplicação será desenvolvida segundo um processo iterativo de desenho centrado no utilizador. Assim, depois de definida uma estratégia inicial, iteram-se as etapas de análise e especificação de requisitos, desenho conceptual e avaliação, até se encontrar o resultado final. Devido a restrições geográficas, primeiramente foram realizadas três sessões de testes em Portugal, participando utilizadores com idades compreendidas entre os seis e os doze anos, com diferentes níveis de conhecimento da língua inglesa bem como diferentes tipos de contacto com dispositivos móveis. Os resultados revelam eficácia nas técnicas de *gamification* aplicadas à literacia de modo a oferecer uma experiência cativante ao utilizador e ainda resultados prometedores sobre as interfaces adaptáveis.

É expectável que esta ferramenta contribua para a melhoria da taxa de literacia na África Subsariana, permitindo a professores e voluntários diagnosticarem e acompanharem facilmente o grau de literacia nessas regiões. Espera-se também despertar nos membros das comunidades o espírito de entreajuda para o combate ao baixo nível de literacia existente nas suas áreas de residência.

Palavras chave: sistemas de avaliação de literacia, ludificação na literacia, reconhecimento de comportamento do utilizador, desenho centrado no utilizador, interfaces adaptativas, comunidades subsarianas

Acknowledgements

I remember the day when computers came into my life. I was only five years old and my father was so excited about that new thing he bought, that for a moment it seemed like we were two kids in that room, and not just me. A few days later, and after I kept insisting, my mother, smiling at me, told me to be careful and then allowed me to turn on the new machine of our home. Perhaps they don't remember that anymore, but those moments still make me smile today.

I started with this little story to thank my mother and my father, who have always supported me and believed in me. They have taught me and my sister the most important values of the life, as well as they know. Thank you Ana, as well, for being my little sister who I admire very much. Thank you, my beloved family, for always standing by my side.

Additionally, I would like to thank Fraunhofer Portugal for providing their facilities and resources for the development of this project and my fellow colleagues for having received me and helping me so well during this semester. Specially, I would like to express my gratitude to my dear friend Eduardo Barbosa, who listened to both my joy and frustration for the entire semester and who inspired me and always helped me with his experience and good will.

Last but not least, I'd like to sincerely thank my supervisors, professor José Luís Cabral Moura Borges and Joana Lobo for their effective guidance, patience and readiness which were essential in getting this dissertation done.

Marcos

"He who does not know can become knowledgeable from learning;
he who thinks one knows and ceases to continue to learn will stagnate"

Akan saying

Contents

List of Figures

List of Tables

Abbreviations

CRUD	Create, Read, Update and Delete
EGRA	Early Grade Reading Assessment
EGWA	Early Grade Writing Assessment
HCI	Human-Computer Interaction
ICT	Information and Communications Technologies
JSON	JavaScript Object Notation
LISTEN	Literacy Innovation that Speech Technology Enables
MILLEE	Mobile and Immersive Learning for Literacy in Emerging Economies
SSA	Sub-Saharan Africa
SUS	System Usability Scale
UCD	User-Centred Design
UNESCO	The United Nations Educational, Scientific and Cultural Organization
VPWA	Volunteer Partnerships for West Africa
XML	Extensible Markup Language;

Chapter 1

Introduction

Sub-Saharan Africa (SSA) is the geographic region of the African continent that is located south of the Sahara Desert. It is right there where the lowest youth and adult literacy rates in the whole world resides. A literate person is defined by The United Nations Educational, Scientific and Cultural Organization (UNESCO) as someone who is capable of reading and writing, with understanding, a short simple statement about her or his everyday life [SU10]. These countries present extremely low school attendance and often there is a limited access to books in some rural areas. The low specialized human resources in education is also notorious, so much that this kind of tasks often belongs to volunteers.

At the same time, the mobile market in this region is revealing a tremendous growing rate, as the mobile penetration rate of unique mobile subscribers is increasing. In fact, it is expected that by 2020 over 50% of all the connections in SSA will be done by smartphones [GSM14].

Thus, an opportunity to combat a severe problem as illiteracy emerges in the SSA countries, by making use of the characteristics and potentialities of the Information and Communications Technologies (ICT), namely mobile phones.

This dissertation provides a study whose main aspiration is to cover that situation and it presents the specification and implementation steps of a proposed mobile solution named LiteracyTracker.

1.1 Context

This dissertation theme was proposed by the *Fraunhofer Portugal Research Center for Assistive Information and Communication Solutions* (Fraunhofer Portugal AICOS), a non-profit private association founded by *Fraunhofer-Gesellschaft*, which is an organization for applied research in Europe. Fraunhofer Portugal AICOS aims to explore technological innovations oriented towards economic growth, social well-being and the improvement of the quality of life of its end-users.

It also promotes and coordinates cooperation between its research centers, external research institutions and industry partners, with the objective of bringing the product of applied research to private and public enterprises and to society in general [Por].

The resulting work of this dissertation should relate to what is being developed in the *ACP Street Libraries*, a project within Fraunhofer Portugal AICOS and other institutional partners that can make use of this technology. One of these partners is the *Volunteer Partnerships for West Africa* (VPWA), a non-governmental organization based in Ghana, a SSA country. The *ACP Street Libraries* is one of the VPWA projects that aims to promote children education and reduce illiteracy in the Ghanaian communities. For that reason, VPWA needs to assess the literacy levels to observe the impact of their project and keep track of the type of educational material that should be provided to each child [Por15].

Even though a prototype was tested in Portugal at an initial stage, given the geographical constraints, one last and improved version of LiteracyTracker will be tested on field by the VPWA staff members. To meet the target audience's needs, three testing sessions were conducted, participating users aged between seven and twelve years old, having distinct levels of knowledge of the English language and different experiences with mobile devices.

1.2 Motivation and Objectives

In the most isolated populations of the SSA countries there is not an easy and free way to check and assess their literacy levels. The assessment process is often based on pen-and-paper tests and, as was stated before, not conducted by specialized people because this kind of human resources is lacking. This is the case of VPWA, for instance, where the diagnostic tests are conducted by one of the organization's volunteer members and it is confined to a standard A4 sized paper. Adding to this, SSA populations show a lack of awareness of the relevance and impact that literacy and learning process can both take in their lives.

The research conducted showed some studies and tools to solve the stated problematic, but they don't meet all the requirements this dissertation proposes to. Most of the applications presented themselves merely as reading tutors, not evaluating speech or even writing capabilities of the people under testing. Adding to this, the research reveals that most of them do not contemplate any kind of engagement or fun for children.

For this situation, mobile phones may take an important role in tracking literacy and also in motivating people to learn how to read and write. The user engagement in reading and learning is essential and it can be improved through the gamification of literacy. Simultaneously, specific behaviours during the literacy assessment can be identified and analysed, helping to define user profiles that will help to adapt interfaces for each user and mitigating the need for professionalized people to conduct the tests.

Having this in perspective, the goal of this dissertation is the specification and implementation of a gamified mobile application that will monitor the literacy levels in the SSA countries. Moreover it should be suitable for the socio-cultural context proposed. The solution should be achieved

by:

- autonomously assessing user's literacy level through engaging tests;
- tracking user's literacy progress over the time;
- correlating user's literacy level with his behaviour while interacting with the application;
- modifying the application interface and usability according to the literacy profile and behaviour of the user;

1.3 Dissertation Structure

Besides this introductory chapter, this dissertation contains five more main chapters. Chapter 2 offers an overview of the state of the art of three important aspects inherent to the presented problematic and objectives: literacy assessment systems, gamification and human-computer interaction (HCI) guidelines for illiterate users and SSA as well. Chapter 3 and Chapter 4 describes the specification of the solution and the implementation of a prototype. Decisions made are explained as well. Chapter 5 shows the user testing sessions and the respective results. It analyses the three conducted sessions of tests, the followed procedures and describes the participants. The results obtained are presented as well. Finally, on Chapter 6, some conclusions are made.

Chapter 2

State Of The Art

Chapter 2 provides a review of work in three areas of study considered important to better understand how to develop and achieve the proposed objectives of this dissertation. Thus, the following sections are divided into three main topics: Literacy Assessment Systems, Gamification and HCI for sub-Saharan Africa.

2.1 Literacy Assessment Systems

There have been many efforts to measure literacy acquisition over the past decades, so a wide variety of tools for assessment exist. Relatively to developing countries, evaluations of literacy programs are numerous. However, this is a critical situation. In some societies, for example, literacy may be simply inferred from school attendance: those with a specified minimum number of years of formal public schooling are assumed to be literate [Wag03]. Measures like these, however, may be unsatisfactory, since it is not uncommon in some poorer countries to have attended school without having acquired sustainable literacy skills [Ter]. In other countries, literacy rates are calculated by just asking people a simple question "Can you read and write?". Based on this information gathered in a household survey or census, the result answer will define if the person in question is then literate or not. [Wag03, Ter, fS]. Obviously, questions like these generally lead to overestimates of literacy rates, due to a variety of reasons. These few examples illustrate the diversity of difficulties in producing comparisons and retrieving conclusions about the state of literacy globally [Ter].

Assessment refers to one or more methods for judging the actual performance of literacy or other cognitive skills [Wag03]. There are several known traditional techniques for assessing literacy as the oral reading fluency, the written spelling or others based on pen-and-paper tests, but there are also some diverse technological frameworks.

One of the most recognized tools designed to measure basic foundation skills for literacy acquisition is The Early Grade Reading Assessment (EGRA). EGRA is an oral test that examines a

person's ability to perform fundamental pre-reading and reading skills such as evaluating knowledge of vocabulary, phonemic awareness, word comprehension among others. The person takes about fifteen minutes to perform the test [Deb11, Gov08].

Even knowing that writing skills are an integral part of basic literacy skills, tools for assessing such skills have been less developed, compared to those for reading, on which significant work has been done [Une15]. In this particular context, UNESCO ordered a study on formative assessment with the final intent of developing an instrument for early grade writing assessment (EGWA). EGWA comprises a list of tasks such as writing the alphabet in order from memory, writing words and sentences, and copying or writing sentences from dictation [UNE12].

In [Wag03], three parameters are suggested for consideration when new tools for literacy assessment need to be developed for developing countries. In short, this kind of tools should be:

- **Smaller**: assessment methods should be robust enough to answer important policy questions at the national and local levels. *Smaller* here generally refers to: first, the number of countries included in such studies may be only one, and in the case of big countries it may be at the state level; second, the population assessed, as well as the number of items utilized in assessment instruments should be just big enough to answer the relevant questions.

- **Quicker**: literacy assessments need to be completed in a way that can affect policy and spending in the lifetime of current ministerial appointments of the country.

- **Cheaper**: these countries cannot afford high costs. The higher the cost involved in the application, the more difficult it is to get a permission to participate in such an exercise.

2.1.1 Application Examples

In this subsection, a list of some literacy acquisition application examples is presented.

- **MILLEE**: MILLEE (Mobile and Immersive Learning for Literacy in Emerging Economies) aims at developing cellphone applications that enable children in the villages and slums, at times and places that are more convenient than schools, to acquire language and literacy in game-based environments. MILLEE was tested in multiple communities in north and south India. The final result is not only the game, but a suite of tools and methods for adapting and extending them for local community use [Tew15, KAK+08].

- **LISTEN**: Project LISTEN (Literacy Innovation that Speech Technology Enables) is an automated computer based *Reading Tutor* that displays stories on a computer screen, and listens to children read aloud. This project was developed at Carnegie Mellon University with the goal of improving literacy. It analyses the child's oral reading and intervenes when it notices the reader made a mistake, or encountered some difficulty. It gives spoken and graphical feedback based on expert reading teachers. It was already tested in schools in the United States, Canada, Ghana, and India [tL15, Deb11].

- **Tangerine**: Tangerine is a digital assessment interface designed for use on mobile computers, tablet computers and smartphones. One of its primary uses is to enable the recording of students' answers in oral EGRA, and interview responses from students, teachers and principals on the home and school contexts. It can be used for the standard EGRA approach, or it can be customized for other types of surveys such as math diagnostics or school information surveys [Tan, Deb11].

- **GraphoGame**: GraphoGame is an evidence-based learning game that helps children to learn how to read. Initially, by playing the game, children learn first the basic letters and their sounds and then through a series of levels, they gradually move on to short and increasingly longer words. It has an adaptive interface, since it dynamically adapts the difficulty level to the child's profile [Inf, RKTL14]. Several versions of GraphoGame have been developed for different user groups, such as immigrant children who are learning how to read in Finnish and for native-speaking children who need remedial training in reading fluency [RKTL14].

2.2 Gamification

In order to give a sense of engagement to applications or activities that don't seem so interesting and motivating to the user, a new concept appears. This section aims to present a review of this design approach called *Gamification*.

2.2.1 Gamification Definition

The concept and the importance of gamification are somehow controversial. In the literature there is no agreement when it comes to the definition of gamification and even its usefulness is questioned [Rug13]. Nevertheless, through the different found definitions [ZC08, Det11], one widely used explanation suggests gamification as the "use of game design elements in non-game contexts" [Det11]. Decomposing this definition, three main aspects appear: game elements, game-design techniques, and non-game contexts.

In short, a game manifests itself as an integrated experience, but it is built from many smaller pieces. Those are called game elements [WH12]. In [Sch08], game design is defined as "the act of deciding what a game should be". In other words, gamification uses game design techniques in order to choose which game elements are useful in a given context to grant an engaging experience. Here resides another important aspect of the given definition: gamification operates in non-game contexts, which means that its challenge is to take elements that normally operate within the game universe and apply them in the real world [WH12].

Gamification is easily confused with other related terms such as *games*, *playful design* or even with more distinctive ones like *game theory* [Det11, WH12, Wu11] . Still, with the *serious games* concept this problem is particularly recurrent. In fact, the main purpose of both phenomenons is other than enjoyment. However, in contrast to serious games and entertainment game-based

approaches, the idea of gamification is to build a system which contains several design elements from games without being a full-fledged game [Det11, BAG14].

2.2.2 Game Elements

The number of game elements that can be adopted by gamification is extensive. Those elements may be divided into three categories that are relevant to gamification: *Dynamics*, *Mechanics*, and *Components*. This is a hierarchical model that is in the form of a pyramid. The pyramid model, represented in Figure 2.1, contains the three categories organized in decreasing order of abstraction in a way that each mechanic is connected to one or more dynamics and each component is connected to one or more mechanics or dynamics [WH12].

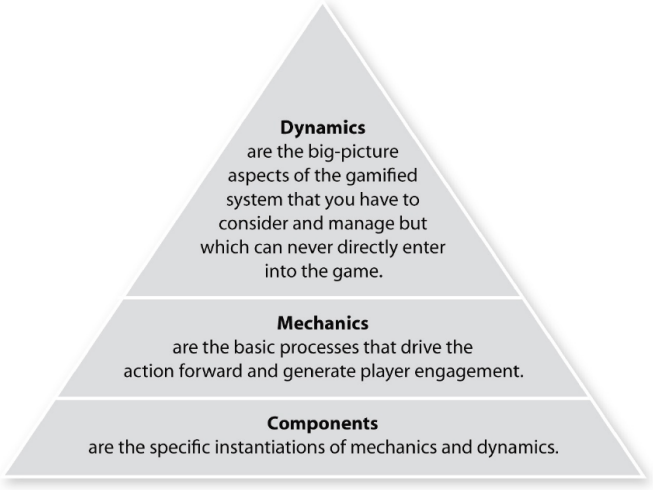


Figure 2.1: The game element hierarchy, model developed by Prof. Kevin Werbach [WH12]

.

Dynamics are the top level of thinking with mechanics coming as second and components as third. The higher level directs the lower level and several lower level components can be used to achieve a higher level goal [Kuu13]. The integration process of these three categories of elements is the crucial task of gamification design process. It is important to retain that no gamification activity requires using all elements. The designer need to ensure that the chosen elements match the requirements of the context and that the application will benefit from all of them [WH12]. In short, the best gamified experience is not the one which uses the most elements, but the one which uses them most effectively.

In the next subsections each category is individually described.

2.2.2.1 Dynamics

At the top of the pyramid are game's dynamics. Game dynamics can be defined as the "temporal evolution and patterns of both the game and the players that make the game (or any gamified

activity) more enjoyable" [Wu11]. These are the most abstract game elements. Players feel their effects but do not engage with them directly [WH12]. Table 2.1 shows the most significant high level elements in a game or gamified system.

Table 2.1: Dynamics (adapted from [WH12] and [Kuu13])

Element	**Description**
Constraints	Limitations of the game that affect the players' freedom in order to create meaningful choices and interesting problems
Emotions	How the player feels: curious, competitive, frustrated, happy
Narrative	The structure that makes the game into a cohesive whole. A consistent, ongoing storyline with a purpose, not just a collection of good ideas
Progression	The improvement of the player in the game, instead of giving the idea of doing the same action repeatedly
Relationships	The interaction players have with other players (friends, teammates or opponents)

A well designed gaming dynamic may led players to a feeling of accomplishment. On the other hand, poor gaming dynamics tend to lose players during the experience, either due to boredom or too many complex challenges [Wu11].

2.2.2.2 Mechanics

In the middle of the pyramid are game's mechanics. Mechanics can be considered as tools that help to determine how to drive the action forward and engage the players into the game [WH12]. Basically, they stimulate the behaviour of the player in the desirable direction. In Table 2.2 we can identify ten central game mechanics of this model:

Table 2.2: Mechanics (adapted from [WH12] and [Kuu13])

Element	**Description**
Challenges	The game sets an objective for the player to reach, like puzzles or other tasks that require effort to solve
Chance	Some elements are randomized which creates a sense of surprise and uncertainty
Competition	One player or group wins, and the other loses
Cooperation	Players must work together to achieve a shared goal
Feedback	Information about how the player is doing in the game
Resource Acquisition	The player can collect items that help him reach the goal of the game
Rewards	Benefits for some action or achievement
Transactions	Trading between players, directly or through intermediaries. Transactions can also happen with non player characters
Turns	Every player in the game has their own time and opportunity to play
Win States	Objectives that makes one player or group the winner. Draw and loss states are related concepts

Each one of these elements is a way of achieving one or more of the dynamics described [WH12]. For example, with feedback(mechanic element) the player can get a feeling of progression(dynamic element).

2.2.2.3 Components

At the bottom of the pyramid, with the lowest level of abstraction, are the components. Therefore, these are the most concrete elements which may be used and seen in the interface of the game. Just as each mechanic ties to one or more dynamics, each component ties to one or more higher-level elements [WH12]. An easy example: points and badges(components) can be part of rewards(mechanic element). Table 2.3 shows fifteen components which the hierarchical model considers most important.

Table 2.3: Components (adapted from [WH12])

Element	**Description**
Achievements	Defined objectives
Avatars	Visual representations of a player's character
Badges	Visual representations of achievements
Boss Fights	Especially hard challenges at the culmination of a level
Collections	Sets of items or badges to accumulate
Combat	A defined battle, typically short-lived
Content Unlocking	Aspects available only when players reach objectives
Gifting	Opportunities to share resources with others
Leaderboards	Visual displays of player progression and achievement
Levels	Defined steps in player progression
Points	Numerical representations of game progression
Quests	Predefined challenges with objectives and rewards
Social Graphs	Representation of players' social network within the game
Teams	Defined groups of players working together for a common goal
Virtual Goods	Game assets with perceived or real-money value

2.2.3 Psychology Behind Gamification

Gamification encompasses a number of psychological concepts, especially regarding motivation, behaviour, and personality types [Bad15b], which are relevant to better understand why people play. A few theories are known in that field of study. More specifically the *Four types of fun* [Laz04], *Self-Determination Theory (SDT)* [Rya09] and *behaviourism* are often discussed in games and gamification literature.

The *Four types of fun* theory indicates that there are four dimensions of fun in a game:

- **Hard fun**: where a player is trying to win some form of competition, which is fun because of the pleasure of overcoming it.
- **Easy fun**: where a player is focused on exploring the system. It is a casual enjoyment.

- **Altered states of fun**: in which the game changes the way the player feels. It's the enjoyment of trying out new personas and new experiences.

- **The people factor**: also known as "social fun", is the category of fun in which the player interacts with other players, even if competing [ZC08, WH12].

The kinds of fun a gamified system should provide will depend on the context. One good way to know is to build it, test it and refine it through a rigorous design process [WH12].

Another aspect for understanding player motivations is by questioning where motivations come from [ZC08]. Motivation is often divided into two sub-types: *intrinsic* and *extrinsic*.

- **Intrinsic motivation**: refers to an innate drive to perform some action, deriving from the core human self [ZC08]. For example, curiosity is motivated by the need to learn. The SDT is a theory that reiterates the importance of this sub-type of motivation. People are intrinsically motivated if they engage in an activity without any hope of an external reward. According to SDT, such activities evoke feelings of *competence*, *autonomy*, and *relatedness* [WH12] which are labelled as the three core intrinsic motivations.

  Gamification uses the three intrinsic motivators to generate powerful results: game elements like levels and the accumulation of points can all be markers of *competence*. Also, giving choices to the players as they progress increases the desire for *autonomy*. Finally, the possibility of sharing achievements or badges with friends, is an example that respond to the human need for *relatedness* [WH12, Kuu13].

- **Extrinsic motivations**: On the contrary, extrinsic motivations are driven mostly by external aspects of the real world, pushing the player to do something for a reason other than its own sake. For example, the desire to make money [ZC08, WH12]. With the SDT it is possible to see what motivates people but there is still a problem of activating behaviour [Kuu13]. Behaviourist thinking, like the *Fogg Behaviour Model* [Fog09] suggests that extrinsic motivation is the way to encourage people to perform actions. A reward or punishment, methodically applied, would condition and reinforce responses in anticipation of more rewards or punishments [WH12].

When it comes to intrinsic versus extrinsic motivation in a gamified or motivational design context, there are different points of view [ZC08]. Intrinsic motivation is important for a long lasting relationship between a system and its users. On the other side, extrinsic motivation can motivate people to complete boring tasks, which may then increase intrinsic motivation. Despite that fact, in many cases such as gamification of education, it is important to be careful since using too powerful extrinsic motivation may diminish intrinsic motivation [Hag12].

The point here is that, supported by the knowledge of these emotional instruments, any potential designer of a gamified application should be able to decide which motivators can most effectively succeed in his system.

2.2.4 Player Types

Understanding the personality type of the players is also important in gamification. Individual differences may need different approaches to increase engagement and desirable behaviour. In gamification, these are thought of as *player types* [Bad15b]. To better classify the types of players in a gamified system, the *Bartle Model of Player Types* is commonly used [Bar96]. As is presented in Figure 2.2, the four types of players suggested by the model are: *Achievers*, *Explorers*, *Socializers*, and *Killers*.

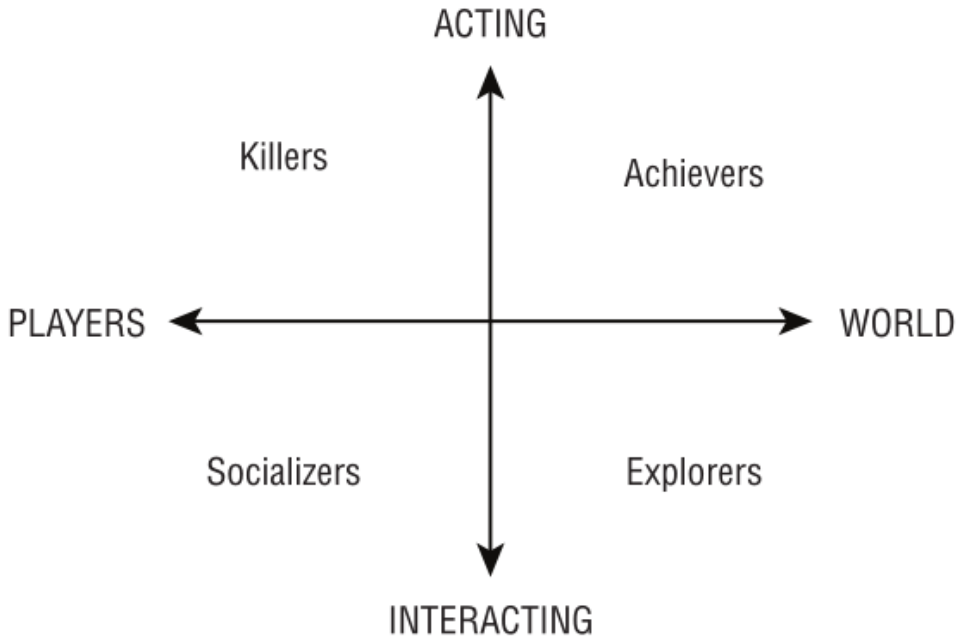


Figure 2.2: The four types of players proposed in the Bartle Model of Player Types [ZC08]

- **Achievers**: Achievers are players who act on the world and are driven to achieve concrete measurements like points, levelling up or earning badges to succeed in a game [ZC08, WH12, Gal11]. They like the sense that they can become more powerful, making progress or advancing certain concrete goals [Gal11]. Designing exclusively for this kind of player raises the problem that it is difficult to develop a system where everyone can win and achieve. Also, to achievers, losing the game is likely to cause them to lose interest in playing it [ZC08].

- **Explorers**:Explorers are players who like to interact with the game's world. Generally, this kind of player likes to find secrets, hidden levels, learn about the story of the game or about content that nobody else knows [ZC08, Gal11]. They often feel restricted when a game expects them to move on within a certain time, because that does not allow them to look around at their own pace [Bad15b].

- **Socializers**: Socializers are players who use games as a way to enjoy interact with others, and on some occasions, player-controlled characters with personality. It is more important for them to interact with others than to achieve something, however, this does not mean that they do not care about the game or winning. [ZC08, Gal11]. Thanks to the social media explosion, this type of players is the one that has grown the most [Hag12].

- **Killers**: Killers are players who act on other players. For them, winning against the environment of the game (e.g.: player-controlled characters) and defeating other players are different things [Gal11]. They are similar to Achievers in their desire to win but, unlike them, winning is not enough for Killers. They must win and someone else must lose and because of that they like to appear at the top of leaderboards [ZC08, Gal11]. Killers want to impose their will on others, typically by vanquishing them [WH12].

It is necessary to note that people are not exclusively one or another of the four player types. A player may even meet characteristics of all types at the same time. The proportions vary in different environments, and a player's primary motivation can shift over time. Even the killers may be your friends if they galvanize everyone else in a positive way, for example. Thus, taking the perspective of a game designer, it is a challenge to see how people are motivated to play and interact in a gamified system [ZC08, WH12].

2.2.5 Applications

In previous section 2.2.1, it became clarified that gamification operates in non-game contexts. This basically means that gamification can be used in many different areas such as health, productivity, environment,education, marketing [Kuu13], enterprise [WH12, vdB14], tourism [Gon13], among many others [Bad15a].

In what concerns education and learning systems, several gamified applications exist. Some of the most important of which are described below. Through clever use of gamification they have succeeded in making education a fun task and received much publicity thanks to their use of gamification [Hag12]. Adding to this, all of them count with millions of registered users.

2.2.5.1 Khan Academy

Khan Academy is a non-profit organization with the mission to offer free high education to anyone, anywhere. It offers practice exercises, instructional videos, and a personalized learning dashboard that empower learners to study at their own pace in and outside of the classroom. This is an example of a web platform where it is possible to learn learn maths, science, biology, chemistry, computer programming, history, art history, economics, and more. [ohtkotYcla15]. Khan Academy has many gamification traces in order to motivate those who seek for further education, such as cumulative rewards, big list and kinds of badges, possibility to unlock new classes and learn new skills, instant feedback, progression bars, among others.

2.2.5.2 Codecademy

Codecademy is another website offering free education to everyone [Cod15]. It teaches how to code in different programming languages by *gamifying* the learning experience. To engage players, the site presents elements such as feedback, points and badges for completing tasks in lessons. But the social component is relevant too: there are networking options to allow users to code with their friends and monitor each others' progress, keeping the essential motivation levels high.

2.2.5.3 Duolingo

Duolingo offers free language instruction where users translate documents and vote on the accuracy of those translations. It is a web and mobile platform.

As Duolingo proclaims: "gamification poured into every lesson" [Duo15]. That is probably the best way to introduce the relation between Duolingo and gamification. It presents lots of game elements since the beginning of the experience. Each lesson includes a variety of speaking, listening, translation, and multiple choice challenges. The player can earn points for correct answers, race against the clock, and level up. The concept of unlocking courses is also present. When missing a challenge, Duolingo quickly shows how the player can improve. It also tries to motivate the player to stay on track by recording how many days in a row the player spends learning a language [Duo15]. Another important aspect is that the system is adaptive: it tracks each completed lesson, translation, test, and practice session to provide feedback to the players and plan future lessons and translation assignments to better address their needs.

A study conducted to evaluate the effectiveness of Duolingo concluded that the vast majority of the participants, which were non Hispanic and with no knowledge of Spanish, liked the product and most of them succeeded in improving their knowledge in that same language. It also concluded that the main factor for higher effectiveness was the motivation of the participants [VG12].

2.3 HCI Guidelines

Reviewing the literature of human-computer interaction, it is relatively easier to find different HCI studies and some guidelines for developing countries as a whole. Even research on improving literacy and educational levels in developing countries is a common topic in the HCI community [TLW+10, FBT09, KRD+07]. However, when considering the scope of this dissertation, Sub-Saharan Africa encloses many political, cultural and social backgrounds, implying different user profiles, needs, skills and habits. So, it is important to reduce the scope of research in order to better analyse all those variables and possible obstacles, focusing then on a minor sample.

Nevertheless, it becomes clear that few HCI guidelines exist for the context of SSA. In fact, that reality should not be taken as a big surprise since many of those who have not benefited from the technological advancements are populations of SSA countries [TS].

At the same time, literature contains many studies about illiterate people and first-time users of different kinds of technologies. But, again, in these cases the main focus of the research is not related to the SSA context.

The following Subsection 2.3.1 gives a general image about the average SSA user, it exposes the technological barriers of the SSA context as well as it lists guidelines to apply in future technological solutions. Next, in Subsection 2.3.2 it is possible to observe a gathered set of guidelines for illiterate users.

2.3.1 HCI Guidelines for sub-Saharan Africa

2.3.1.1 Understanding the Average SSA User

Understanding the average user in sub-Saharan Africa is crucial to define guidelines for the expansion of technology in this region [TS].

The following are some of the characteristics and conditions that may help define the average SSA user. That definition was based on a model often used in HCI to help define an average user called *Hofstede's five dimensions of culture* [Des01].

- **Authority**: most SSA countries reveal a strict hierarchical view of authority. People on the top of the hierarchy are often significantly better paid. Instructions are rigorously followed without inquiring. Parents teach obedience, and expect respect. Teachers possess wisdom and are automatically esteemed [TS, Des01].

- **Collectivism**: In SSA there is a tendency to have a collective way of thinking, emphasizing the community. Collectivist cultures value training, physical conditions, skills, and the intrinsic rewards of mastery. In general, the family values harmony are over all else. The Government may enforce policies that favour the group over the individual and investigations into a person's private life is tolerated. Other policies may include controlling the economy and the press [TS, Des01].

- **Gender Roles**: in relation to gender roles in society, this region is neither a highly masculine nor a highly feminine culture. There is some blurring of roles of men and women [TS, Des01].

- **Coping with the unknown**: some cultures have difficulties in cope with uncertainty. This can influence values regarding everything from formality, punctuality or tolerance for ambiguity. In SSA there is some ability to cope with the unknown [TS, Des01].

- **Short term thinking**: the sub-Saharan area is more inclined to short term thinking. Part of that can be associated to the instability of the area, the fact of not being able to make plans for the long term due to constant political and economic changes [TS, Des01].

2.3.1.2 Heuristics and Guidelines

Literature shows that at least one relevant study was conducted in order to evaluate the technological paradigm in SSA, proposing new HCI principles to foster technological advancement in this particular region [TS]. It discusses the different kind of barriers that prevent the existence of an information society in SSA, describes the region's culture and the general role that HCI can take in these environments.

Table 2.4 presents an overview of the detected barriers and related guidelines. Traditional HCI heuristics and guidelines were reviewed and modifications were suggested to make them compatible with the needs of the average user in SSA.

Obstacles	**Existing Guidelines**	**Modification**	**New Guideline**
1.First-time users /Novices	1.Offer error prevention and simple error handling; 2.Make components visible and obvious; 3.Permit easy reversal of actions;	1.Emphasize error avoidance handling; 2.Components and applications named to convey purpose; 3.Ensure users know consequences of actions;	1.Applications should assist user in error recovery;
1.Recent conflicts	1.Simplify the structure of tasks;	1.Goal achieved quickly with minimal steps;	1.Allow various color palettes and images to aid memory;
1.Political Instability; 2.Differences in status of people;			1.Store sensitive data in neutral and secure locations; 2.Use passwords and security features to create separation;
1.Group success more valued than individuals;			1.Create separations that are based on the group's status not the individual's;
1.Blurred gender roles;			1.Incorporate unifying values;avoid gender distinctions;

1.Different languages; 2.Pre-conceived notions(non-local designers);	1.Design for cultural diversity; 2.Use *affordances*;	1.Use *defamiliarization* rely less on *affordances*;	1.Create adaptive applications to handle written and oral communication in multiple languages; 2.Use locally understood and accepted terms and images
1.High illiteracy rate;	1.Offer informative feedback;		1.Applications should "speak" to users to convey options and errors;
1.Inconsistent power source;			1.Auto-save data frequently to non-volatile state; 2.Use innovative power sources like wind up cranks;

Table 2.4: Barriers and proposed HCI guidelines for the SSA context, adapted from [TS].

2.3.2 HCI Guidelines for Illiterate Users

The literature has an extended list of well known studies searching for the best HCI guidelines for illiterate users. This subsection summarizes the gathered guidelines and conclusion of different studies which are expected to instruct designers and developers to achieve this type of user's needs. Semi-literate users are also taken into consideration, since many of these guidelines are also oriented to them.

2.3.2.1 Interface Design

Text

- text-based user interfaces can cause anxiety on low literacy users [MPT07].
- illiterate and semi-literate users respond differently to text in interfaces: for fully illiterate users, minimal text should appear in the interface [FBT09]. However, according to [Gav12], treating every user as illiterate has a negative impact on semi-literate users.
- when text is extended with audio or graphics it reinforces semi-literates' reading skills [FBT09, Gav12].

- augmenting text with audio or graphics also motivates learning and improves literacy [Hue08, JWKS08].
- using numbers in the interface might be plausible since many illiterates and semi-literates are numerate [MPT05, MPT07].
- the interface should allow full operation of the device without requiring users to read [Hue08].
- even when users cannot fully read the text displayed, using their local language gives them a greater sense of familiarity and ownership [PGC03].

Graphics

- semi-abstract, realistic, cartoon representations are preferred and understood more accurately by illiterate and semi-literate users than abstract simple graphics [MPT05].
- hand-drawn cartoons are better understood than photo-realistic representations [MPT07].
- icons are better understood when labelled by textual descriptions [Chi05, LG08].
- an icon should never be used to express meaning on its own, so a form of visual or audio text captioning should always be available [Hue08].
- users who are unable to read text can be assisted by the aid of pictures explaining the text [Gav12].
- some symbols must be avoided when designing for some cultures. Effective icons should tap into the culture, referencing concepts in their lives. Icons should be memorable, nameable, and concrete so that users can discuss them with each other [Hue08].
- meaningful shapes should be used instead of abstract ones (a star instead of a triangle, for example) [JWKS08].
- even if the name of the icon is not described, users should be able to assign them a name [JWKS08].

Audio and Voice

- in [MPT05], some illiterate and semi-literate users were excited when they heard a computer producing voice sounds in their native language.
- all screens should have a ”help” icon that provides voice instructions and aids the user to navigate through the application [MPT05].
- continuously available help is decisive to illiterate users. An audio assistant, for example, is an option [PMTB08].
- step by step help instructions have been shown to be more effective and preferred by low-literacy users compared to the help button [LG08].

- advising the users when they should speak is one of the main requirements for speech interfaces, accordingly to [KRT+12].
- congratulatory audio messages after the user completes a task seem to produce excitement and encouragement [PMTB08].
- information presented orally needs to be short. Low literate and literate users find it hard to hear long passages of text. Sentence structures should be simplified and difficult words replaced [SPM+09].
- users should be able to interrupt, skip or repeat audio options any time [Hue08].
- high recognition accuracy is a necessary condition for the success of a speech system [SPM+09].

Device Manipulation

- the physical design of the phone shouldn't stigmatize illiterate users, that is to say that it should be similar to other devices in the market [Chi05].
- for novice users, pointing with a finger has been shown to be more intuitive and easier than using a pointer [Gav12].
- multi-function buttons should be avoided because it confuses users [LG08].
- direct access to features through single-click actions is the most effective and has highest recall [LG08].

Navigation and Information Architecture

- linear navigation structures are quicker to understand than hierarchical structures for low-literacy users [MMCT10].
- novice users should be encouraged to explore the interface to overcome their technophobia [Hue08].
- the home screen of applications targeted at illiterate and semi-literate users should be kept simple to avoid intimidating first-time users [MST06].
- offering more than one way to accomplish a task might confuse some users, so the number of possible tasks in each screen should be minimized [PMTB08].
- in many studies, vertical scrollbars were not well understood by illiterate and semi-literate users [PMTB08].
- real-time metaphors should be used to explain foreign concepts. In [MST06], illiterate and semi-literate participants seemed to understand quicker hypertext navigation when they were told to think of the pages as pages in a book.

2.4 Summary

This chapter started by a study of the state of the art of the existent literacy assessment systems, in Section 2.1. Examples were given and it was exposed some existent gaps in this field of study, mainly the fact that the studied tools are not evaluating speech or even writing capabilities of the people, as it would be expected.

Section 2.2 showed that gamification is being used in many different areas, namely learning and education, but gamifying literacy is not so usual. From a scientific perspective, this fact increases the need to analyse the effect of these techniques in people's actions in order to persuade them to improve their literacy levels.

Finally, Section 2.3 gives an overview about the proposed guidelines for designers and developers involved in projects oriented to illiterate users living in the SSA context.

Chapter 3

LiteracyTracker: Specification

This chapter describes the project specification of the main system functionalities an behaviours expected form LiteracyTracker. Section 3.1 describes how the proposed system works. Next, in Section 3.2 both non-functional and functional requirements are identified, as well as the actors and respective roles in the system. The literacy test model used in the system is detailed in Section 3.3 and an explanation of how the system is gamified appears in Section 3.4. In Section 3.5 all the decisions related to HCI are listed. Finally, the adaptive interface mode of the system is analysed in Section 3.6.

3.1 Specification Overview

The proposed solution presented in this dissertation that aims to resolve the problem exposed in Chapter 1 is a mobile application denominated LiteracyTracker.

LiteracyTracker aims to evaluate the literacy levels of SSA communities by using gamification techniques make assessments enjoyable experiences for the users. Its interface is adaptive, which means that it changes according to the needs of the users, by analysing their literacy profile and interaction behaviour.

After the user's registration and authentication, the application submits them to a new literacy test model inspired and specially designed after all the research made and showed in Chapter 2. This test is divided into a set of different exercises, each one with its own proper scope of evaluation. Thus, by starting a new assessment, users face a sequence of levels of exercises that they should complete in order to succeed and advance to the next one.

Some aspects are common to every level, however with distinct values assigned depending on the level in question. They are:

- the number of expected rounds that the user has to give answers;
- the number of given answers that can be wrong;
- the number of points to subtract in case of incorrect answers;

- the number of points to add in case of correct answers;

By the end of each test, and relying on the users performance, a literacy profile is associated to them. At this moment, they can also view their test result in the form of points, earned by the correct given answers.

Over time, there are rewards that may be unlocked when eventually the user reaches predetermined goals. These rewards can be consulted at any time. Additionally, users can also check their latest test score, their total points accumulated during all previous performed tests and , finally, their all-time best score for a single literacy test.

Finally, the application interface changes by analysing some factors like the users' behaviour during the past interactions or their literacy profile.

3.2 Specification Requirements

The following section describes the structure, intended behaviour and overall perspective of the proposed system. This is accomplished by listing both non-functional and functional requirements, as well as a view of the system's actors and their roles.

3.2.1 Non-functional Requirements

The designed system must satisfy the non-functional requirements described below:

- **Accessibility:** The system should be usable by a wide set of people, specifically illiterate or semi-literate people, who will be the target audience.
- **Availability:** Due to the lack of connectivity coverage in specific regions, the system should be available when connectivity is lacking. For that reason, apart from the server side, the mobile application should have local databases and synchronizes the data to the server only when connected to the internet.
- **Backup:** In order to prevent data loss related to the user's literacy levels progression, backups of the system should be made periodically. Besides that, due to contexts as the SSA where the risk of intermittent internet connection is high [TS], mobile devices must backup all the data locally since the connection to the server is not always granted.
- **Extensibility:** The system should be developed in such a way that facilitates the possible insertion of new functionalities as new kinds of exercises and game elements.
- **Performance:** Due to the inherent limitations on the target field [TS], some validations must be made to ensure that the system operates as expected when deployed in the real-world scenario. Using lightweight data-interchange format in communication protocols, such as JSON (JavaScript Object Notation), applying data compression techniques, condensing multiple data requests into batches and profiling the application in scenarios that

simulate the expected conditions in the SSA environment as best as possible, are some of the techniques that must be explored.

- **Scalability:** The system should handle a growing amount of existing resources in a capable manner. Essentially it should be prepared to deal with an increase of content related to the level's exercises (list of words, images, sounds, questions and answers), assessments in different languages and registered users.
- **Security:** The system should have well defined security policies to ensure that the content is only accessed and manipulated by those who should be given that privilege. User privileges and user accounts should ensure that.
- **Usability:** As research showed, due to the nature of the end-users, complex interactions should be avoided when possible. Visual or audio cues, for instance, should be provided, helping users to interact with the application when in doubt on how to perform a specific action. Also, the user should interact with the system without the need of professional assistance.

3.2.2 Functional Requirements

In order to understand the main features of the system, LiteracyTracker should meet the functional requirements listed below:

- Assess literacy of the users with a level system in which levelling up implies an increase of difficulty of the exercises;
- The language of the assessment can be selected in the mobile application, if already downloaded from the server the relative language package;
- Mobile application receives updates from the server: insertion, creation and edition of reward symbols, levels' databases, language packages in which the application should be assessing users, among others;
- The system should be distributed, being the server responsible for gathering the local databases from the client mobile application, establishing backups and updating leaderboard tables with user and community scores;
- The interface mode algorithm can be changed to be improved in the server side and then communicated to the mobile client;
- The mobile application should detect whether the device is connected to the internet and then allow the data synchronization with the server;
- List of badges to unlock with defined objectives;
- Audio and text help should be available;

- The performance of the users during the assessments should be recorded in order to monitor their progress;

3.2.3 Actors and Use Cases

In this section the users of the system, commonly known as actors, and the different roles they play when interacting with the system are presented. LiteracyTracker has three types of actors: the Admin User, the Regular User and the Staff Member.

Admin User

The Admin User only operates on the server side of the system. This actor is responsible for reading, creating, updating or deleting(CRUD) content such as the database of each level and reward symbols. This user may also update the algorithm of the interface mode recommendation. After that, the changes made are communicated to the mobile application as an application update. In Figure 3.1 the Admin User's use cases are presented.

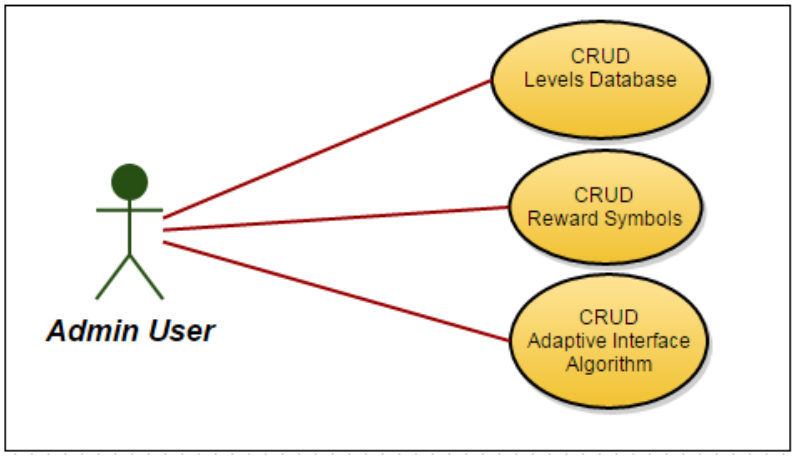


Figure 3.1: Admin User's use cases diagram.

Regular User

The Regular User, which in this dissertation is often simply referred as user or end-user, is the actor of the system that represents the children or adult that will be assessed. Besides the assessment, they can consult their personal results and rewards, as well as the table of leaders of their community. In Figure 3.2, the Regular User's use cases are presented.

Staff Member

The Staff Member is a representation of the person who supervises the child or adult that will be under a literacy assessment. This user does not necessarily need to have a degree or training experience in education. In fact, the main requirement for this kind of user is to be a literate person.

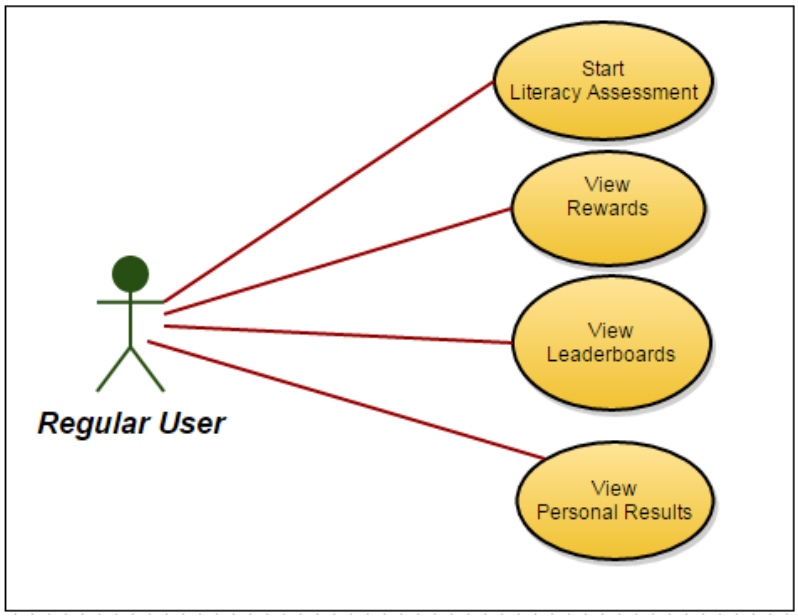


Figure 3.2: Regular User's use cases diagram.

In an organization context, as VPWA, this actor represents a volunteer that conducts the literacy tests on children. The use cases of the Staff Members, shown in Figure 3.3, include registering and authenticating Regular Users and viewing their literacy evolution. Also, their own registration and authentication has to be made, in order to keep record of who was the person that provided the evaluation test to that Regular User. Additionally, the Staff Member may accept, or not, the recommendation given by the system to activate the adaptive interface mode and synchronizes all the data stored locally in the mobile device to the server once a network connection is detected.

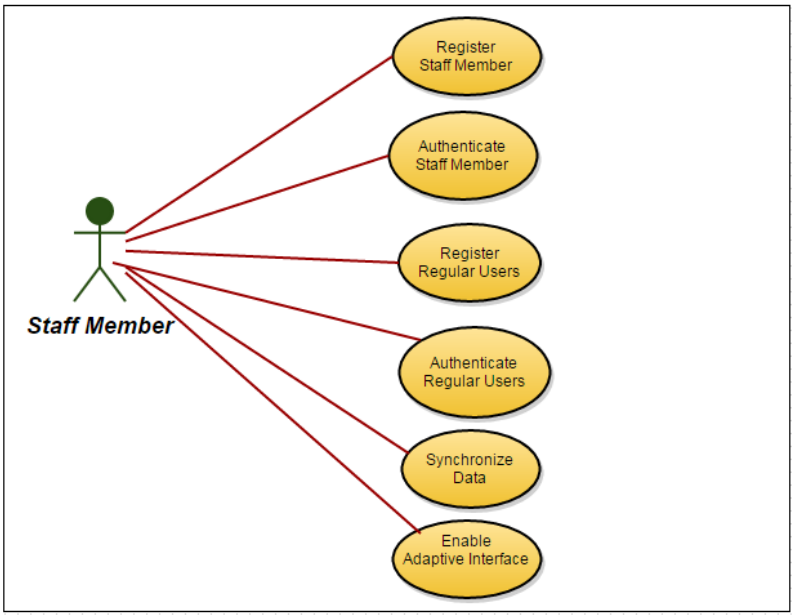


Figure 3.3: Staff Member's use cases diagram.

3.3 Literacy Test Model

Research made reveals that there is not single standard and universally adopted model for literacy assessments. Actually, it exposes the fact that the overwhelming majority of assessments does not cover the two integral parts of the literacy concept: the reading and writing skills. For that reason, it seemed like a natural decision to create a new literacy test model which should fill the announced gaps.

This new model is designed based on some reading and writing tasks of both EGRA and EGWA [Gov08, UNE12] . Adding to this, some of the tasks and also the classification system that dictates the person's literacy level are inspired by the straightforward test from the ACP Street Libraries project, of the VPWA organization. Table 3.1 presents the tasks of the proposed model and a comparison with the above mentioned test's models.

Table 3.1: Comparison between the tasks of the proposed LiteracyTracker's test model and the EGRA , EGWA and ACP Street Libraries models.

Tasks	LiteracyTracker	EGRA	EGWA	Street Libraries
1. Writing the alphabet letters	Yes	-	Yes	-
2. Word recognition (by listening)	Yes	Yes	-	Yes
3. Knowledge of vocabulary	Yes	Yes	-	-
4. Writing words from dictation	Yes	-	Yes	-
5. Comprehension of sentences	Yes	-	-	Yes
6. Writing a sentence from dictation	Yes	Yes	Yes	-
7. Oral reading fluency	Yes	Yes	-	-

The test's procedure is simple. The children, or adults, under testing will be placed into one of eight categories of literacy: from *i)can't read or write* to *viii)can orally read sentences with fluency* . One by one, these categories are associated to each task, by ascending order, and will decide the literacy profile of the examinees at the end of every test. By default, when a new test starts the examinee has the literacy profile *0) can't read or write*. Starting a new task, if the examinees succeed they step into the next task, otherwise the test is over. In the moment they pass from the task n to the task $n+1$, the literacy profile they get is n .

By consulting the Table 3.2, it is possible to understand the main assessment objectives of each case as well as the correspondent literacy profile that may be accomplished by the examinees if they succeed.

The designed literacy model for the system is a critical aspect of the proposed solution. For that reason, and because this dissertation is placed in a different scientific field than linguistics, for instance, the system should be prepared for possible changes that may occur every time new ways of improvement are discovered.

Table 3.2: List of assessment goals and correspondent literacy levels for each task of the proposed test model.

Task	Objectives	Literacy Profile
-	-	*0) can't read or write*
1. Writing the alphabet letters	to assess if the examinee is able to reproduce all the letters of the alphabet	*1) can read and write alphabet letters*
2. Word recognition	to assess if the examinee can read and distinguish words, not just guessing them	*2) can read words*
3. Vocabulary Knowledge	to assess if the examinee knows the basic meaning of words	*3) can read and understand words*
4. Writing words from dictation	to assess whether an examinee is able to write a listened word	*4) can read, understand and write words*
5. Comprehension of sentences	to assess if the examinee answers correctly to literal and inferential questions about a paragraph	*5) can read and understand sentences*
6. Writing a sentence from dictation	to assess whether an examinee is able to write a listened sentence.	*6) can read, understand and write sentences*
7. Oral reading fluency	to assess if the examinee is able to read a narrative or informational text with accuracy and little effort	*7) can orally read sentences with fluency*

3.4 Gamification Model

Gamification is another central point of this presented system because, as research reveals, gamification is not being too much explored in the literacy field. Moreover, if well applied, it can be used to bring engagement and motivation to applications. Thus, in order to define the best model and to avoid jeopardising all of the system with a bad gamification design, the *Six Steps to Gamification* framework developed by Kevin Werbach were followed [WH12]. As the name suggests, the author states that gamification is best implemented in six steps. Deliberately, only the last step focuses on the gamification components. In the other five, the intention of the framework is to oblige the designer to think about how he will map the available techniques onto particular situations [WH12]. Notice that the target users of the system that will be in touch with these elements are Regular Users only.

Following this method, before designing the LiteracyTracker the next questions were answered [Des14, Lel14]:

1. **Define business objectives.** *Why are you gamifying? How do you hope to benefit your business, or achieve some other goal such as motivating people to change their behaviour?*

   The need of gamification in this system arises with the purpose of encouraging illiterate or semi-literate people to learn reading and writing skills. The literacy test will assess the level of literacy of people but it is also important that the activity itself is not a tedious experience for them. On the contrary, it is desirable to motivate people to improve their literacy skills.

So, by gamifying literacy the system will be holding the users attention and will, therefore, contribute to combat the literacy.

2. **Delineate target behaviours.** *What do you want your players to do? And what are the metrics that will allow you to measure them?*

   The desirable target behaviour is that by being assessed once, users will try to improve their last result by performing another test and so on. This result, which is the correspondent literacy profile assigned to users, is the metric measure that will provide feedback about the users' literacy levels progression.

3. **Describe your players.** *Who are the people who will be participating in your gamified activity?*

   The target users are essentially people that are natural from the SSA countries, illiterate or semi-literate, both genders and from a wide range of ages (although the main focus are children). It's hard to predict the personality type of the first-time users of this system, since they can be from so many different backgrounds. Yet, according to the Bartle's model of player types, three of the four types may be frequent: *Achievers*, *Explorers* and *Socializers* [Bar96]. *Achievers* are driven to succeed in a game by achieving concrete measurements like points, levelling up or earn rewards. In the system context, this kind of player will be working hard to be at the top of the community, with the best literacy results. *Explorers* have the need to find something new and love to discover hidden things. So, the curiosity to see what task of the literacy test comes next may let them be intrigued. Also, in the presence of hidden rewards in the system, it may arouse great curiosity in them. Finally, the *Socializers* are the ones who most like to interact with other users then to properly achieve something. However, as previously said, this does not mean that they do not care about winning. After finishing the assessment, sharing their results and accomplishments with the other users might be the most exciting part for them. Also, depending on the assessment space conditions, they will be the kind of player who will enjoy to comment with the other users how they answered in each task.

4. **Devise your activity loops.** *First, describe the kinds of feedback your system will offer the players to encourage further action, and explain how this feedback will work to motivate the players. Second, how if at all will players progress in your system? This includes how the system will get new players engaged, and how it will remain interesting for more experienced players.*

   The tasks of the designed literacy test will be distributed by levels and so the users will have to level up to the next task until the test is finished. Also, for the first time each level is completed with success, a reward should be given to the user. Additionally, for each level, the user has to give more than one answer before levelling up. They receive points for every right answer. In the same way, every wrong answer will deduct points but also reduce the users "lives", i.e, number of wrong answers that they still can give. At the end

of each test, as mentioned before, users can also check their accomplished literacy profile, latest test accumulated points, their total count of points of all performed tests and , finally, their all-time best score from a single literacy test. This score system will allow the users to keep their own progress. All these kinds of feedback and progression loops are expected to increaser the users' engagement and the likelihood that they redo tasks.

5. **Don't forget the fun.** *Consider how your game would function without any extrinsic rewards. Would you say it was fun? Identify which aspects of the game could continue to motivate players to participate even without rewards.*

   The fact that points can increase or decrease each time an answer is given, is one aspect that can motivate users. Furthermore, during all the assessment a character in the form of a pictorial representation of a facial expression will be present on the screen. This character has six different mood states which are references to distinct users actions:

   (a) **neutral mood:** when the user starts a new level;

   (b) **happy mood:** when the user answers correctly;

   (c) **sad mood:** when the user answers incorrectly;

   (d) **awesome mood:** when the user completes the test with success;

   (e) **awful mood:** when the user terminates the test without success;

   (f) **concentrated mood:** when the user is listening instructions from the application;

   Trying to keep the character with good mood is one way to motivate and entertain users. Finally, the ranking score tables between members of each community but also between the participating communities of each country dictate social competition, a factor that often brigs additional motivation to applications.

6. **Deploy the appropriate tools.** *What are some of the game elements involved and what will the experience be like for the players? What specific choices would you make in deploying your system?*

   Table 3.3 shows all the proposed game elements to implement in LiteracyTracker. As it was possible to note, most of them revealed themselves during the thinking process of the previous steps.

Table 3.3: LiteracyTracker's game elements.

Element	**LiteracyTracker Usage**
Achievements	Achievements are presented as defined goals that the user has to accomplish in order to earn rewards. Only performance contingent achievements, which require skill to complete, are present in the system [Kap15].
Avatar	A character in the form of a facial expression of the player is present in the system. This character has six mood states: neutral, happy, sad, awesome, awful and concentrated.
Badges	LiteracyTracker's badges are unlocked whenever the user levels up for the first time, or reaches other proposed challenges. For instance: "complete at least one test without making mistakes". The application only contains skill badges, awarded based on direct user performance and none participation badges, typically earned through a simple act of attendance and no other associated criteria. [Blo15]. Skill badges, may be more easily connected to internal motivation, and thus considered by learners as an intrinsic motivator. If badges are offered for learners who might not excel in the content area of the badges, a negative motivational effect can appear [ASH13].
Content Unlocking	Rewards that need to be unlocked are present on LiteracyTracker.
Leaderboards	LiteracyTracker contains visual displays of individual users progression in the community, ranking by the all-time best score they accomplished. Two leaderboards are in the LiteracyTracker application: the *Community Leaderboard*, where individual progression of the users are shown, and the *Country Leaderbord* where progression of the communities of a specific country are listed.
Levels	Levels are the representation of the defined tasks of the literacy test model. They have associated points, rounds, user lives, rewards and each one is a step in the user progression.
Points	In LiteracyTracker, points are numerical representations of game progression presented in each level. Each level contains points to add for correct answers and points to subtract for incorrect answers. At the end of the test, the total points earned are compared with the previous results and the both community and country leaderboards updated.
Teams	In LiteracyTracker, the teams are represented by the communities. One community is composed by a group of users working together for a common goal: competing with other communities for the best group literacy results of that country.

3.5 HCI Specifications

This section describes the main HCI decisions made, based on the literature review, with regards to some of the major barriers that illiterate and semi-literate SSA users could have to confront. Next, the followed guidelines and decisions made about interface design, device manipulation, navigation architecture and feedback are explained.

3.5.1 Interface Design

When it comes to interface design, LiteracyTracker has minimal text appearing in the interface because, as research showed, it helps low literacy users [FBT09, MPT07]. The interface presents also some audio and graphics elements, mainly in the exercises of the literacy assessment. Due to the audio help tips, the system is also prepared to be operational without much text.

Moreover, the representation of each test level uses the format "Level *n*", where *n* is a number. This decision was made since research revealed that numbers are well accepted by many illiterates and semi-literates who actually are numerate. Describing levels like that, will likewise make unnecessary text disappear, turning the interface even clearer.

When the adaptive interface mode is set to active, the system presents also text and audio in the native language, because, as mentioned before, even when they can't read it, it gives users a sense of familiarity and ownership [PGC03].

Help buttons are always available with voice instructions. The accent of the voice is determined by the voice packages installed on the user's smartphone. However, initially LiteracyTracker checks what voice packages are installed and compares them with the country's language of the users. If the application doesn't find any equivalence, it will emit an alert to the Staff Members to install the required package.

Regarding the graphics, they are used essentially in three occasions: in the avatar, command bar icons, levels exercises and in the reward's symbols. For the first three cases, the graphics are as similar to hand-drawn graphics as possible, in the sense that the meaning has to be well perceived independently of the background contexts of the users [MPT05, MPT07]. Futhermore, all icons are labelled with a description and represented by meaningful shapes instead of abstract ones [JWKS08, Chi05, LG08].

In what concerns to the rewards, they should be represented by cultural symbols of the country in question, whenever possible. There are many different social and cultural contexts in SSA, sometimes even in the same country, so when designing the reward system it is imperative to ensure that the target users will understand the meaning of what they have acquired. This can be worse if the chosen symbols somehow collide with the personal values and beliefs of the users. To conclude, if this requirement fails, a relevant part of the motivation and fun of the system will likely fail too.

3.5.2 Device Manipulation

Regarding the device manipulation, exercises that do not need writing skill are designed to use only simple pointing gestures with the finger, since it was shown to be more intuitive and easier than using a physical pointer [Gav12]. When writing is demanded, the on-screen keyboard appears with a minimalistic layout: auto-correction and suggestions are disabled, as well as other special symbols [Msd15]. The idea is to output only the default alphanumeric keyboard layout and not confuse less experienced users with mobile devices. Multi-function buttons, double-tap interactions and rotation of the screen are disabled as well for the same reason.

3.5.3 Navigation and Information Architecture

LiteracyTracker's navigation structure is as linear as possible. In the literacy assessment, since it is completed step by step, that is easy too implement. Starting from their home page, users can start in the first level and just have to go further until the test is done. During the test, rewards appear in the middle of the levels and in the end of the test all the results are presented to the user in a single page. Nevertheless, to revisit the earned rewards or leaderboards tables whenever they want, a simple hierarchical structure with just one sub-level is needed too. For that reason, all pages have a small indication on the screen with the path that goes from the home page to the current location of the users.

Scroll bars are avoided since research reveals that they are not well inferred by illiterate and semi-literate users [PMTB08].

3.5.4 Feedback

The main feedback is present in the LiteracyTracker essentially when the user answers. Whether the answer is correct or not, the system returns visual and auditory responses to the users. Also, the avatar changes its facial expressions according to different user actions.

3.6 Adaptive Interface Mode

The application has two interface modes: normal and adaptive. By default it starts with the normal interface. Then, at the end of each assessment LiteracyTracker calculates what interface mode is more suitable to the users by analysing a defined number of factors. The metrics used for this purpose can change every time a new formulation reveals to be better than before. In that case, the server will communicate the alteration to the mobile application. The metrics considered in the calculation are listed below:

- age of the user;
- number of times the help features of the application are requested by the user;
- amount of time a user takes to answer questions of each level;

- current literacy profile of the user;

Under these factors, the application gives a recommendation of the more suitable mode for the current user. That is just a suggestion, not an imposition. Thus, as previously stated, it is the responsibility of the Staff Members to decide whether to activate the suggested mode or not, since they have one other metric that the application can't use to decide: human perception.

Once the adaptive interface is enabled, some changes will occur. Mainly, the next time the user starts an assessment, an explanation of the level objectives and the expected actions that the user should take will be shown. It's desirable that the users understand what the interface requires from them. The assessment exercise must be the only challenge they face, so they can't lose time guessing where they have to tap, write or perform any other kind of action. One of the suggested guidelines found in the literature review referred that step by step instructions have been shown to be more effective with low level users [LG08]. Based on that, the explanation of the level goals shows in a step by step process, along with the correspondent audio support, what the next level requires. It may contain pictorial representations of some interactions as tapping one answer box, for example. Additionally, in case the assessment language is not the user's first spoken language, the text and audio support may be presented in the native language of the user.

Finally, increasing buttons size, displaying only the indispensable text to reduce the visual noise and changing audio and visual feedback messages for right and wrong answers accordingly to the user age, are other kinds of interface adaptations that the application performs.

3.7 Summary

In this chapter the specification of the proposed solution was explained. Section 3.1 showed an overview of the way the system works. The functional and non-functional requirements of the system were identified, as well as the actors and their roles, in Section 3.2.

The four main layers of this solution were also detailed. First, in Section 3.3 a new literacy assessment model was presented, based on the literature review. Next, the gamification model of the system is fully detailed in Section 3.4, where the *Six Steps to Gamification* framework was described and followed in order to define the best model for the system. Section 2.3 described the followed HCI guidelines and decisions made about interface design, device manipulation, navigation architecture and feedback were explained. Finally, in Section 3.6 were explained the differences between the normal and the adaptive interface mode of the proposed solution.

Chapter 4

LiteracyTracker: Implementation

This chapter presents one implementation of the proposed solution, with its design and technological decisions.

Firstly, Section 4.1 shows some initial considerations about the implemented system. The requirements of the system are listed in Section 4.2. The remain sections explain the design decisions taken in the developed solution.

4.1 Implementation Considerations

First of all, it should be noted that the developed system is a proof of concept, therefore this is an incomplete implementation whose purpose is to verify that the idealized system presented in the Chapter 3 has the potential of being used in the real-world.

The implemented solution was developed according to an iterative process of User-Centred Design (UCD). In [Nor13], UCD is used to describe design based on the needs of the user, leaving aside what he assumes to be secondary issues like aesthetics. The iterative principle progressively refines the design through evaluation from the early stages of design, enabling the designers and developers to incorporate user feedback until the system reaches an acceptable level of usability [Usa15a].

As already mentioned, one last and improved version of LiteracyTracker will be tested in Ghana by the VPWA. For that reason, all of the developed solution was designed thinking of Ghanaian children as the target users. More than 250 languages and dialects are spoken in Ghana, yet English is the country's official language and also the language used for educational instruction [Gha15]. Thus, the language used for the assessments was English.

In this dissertation, three main components are explored: the literacy assessment, the gamification of literacy and the adaptive interfaces through the user behaviour analysis. In this implementation process, these same components were approached by following the research and specifications made.

One of the main differences between the implemented solution and the specified in the Chapter 3, is related with the system's architecture. The implemented solution does not contemplate a

server because it would take too much effort to develop the server side just to evaluate the system's distributed component. For that reason, this version of LiteracyTracker was implemented with the only intention of to evaluate and refine the three elements referred before. However, it was told that an opportunity to test the solution in Ghana, through the VPWA, could appear earlier than expected, so a server side to allow the backup of the assessments on field would be mandatory. Thus, the decision made was to backup the content directly to any free service of cloud storage. The *Google Drive* service was the chosen one. In respect to the functional requirements that were dependent to the server side of the system, they were implemented locally. The changes of the language of assessment content or updates related to reward symbols and levels' databases can be made in the client side by changing the respective Extensible Markup Language(XML) configuration files.

Regarding the proposed gamification model, in Section 3.4, the leaderboards and the teams were not developed, mostly due to the lack of time. But after the implementation of the other gamification elements, the major concern was to ensure that at least an early version of the adaptive interface mode would be implemented and evaluated too.

4.2 System Requirements

In this Section the system requirements are described in order to correctly run the implemented solution.

- **Operating system:** the mobile application was developed and tested on devices running the Windows Phone 8.1 mobile operating system version, or newer.
- **Speech language package:** LiteracyTracker uses text-to-speech and speech recognition features, which are available in different language packages. For the same language, the user can choose different accents from the desired country and also male or female voices. These features can be selected in the device settings. When the desired language is not installed, it can be downloaded and the operating system will install it. Usually, by default, the device automatically installs the speech package based on the language the user chose to use in the operating system.
- **Data Synchronization:** the device should be connected to the internet in order to synchronize all data. When disconnected, the application data is saved in the isolated storage folder. Isolated storage is a mechanism that ensures that other installed applications will not access the data, keeping it safe. After, whenever the device connects again, this data will be sent to the cloud service.
- **Cloud Account:** the implemented solution uses *Google Drive*, a free service that allows cloud storage and synchronization of files. For that reason, the application will need a *Google* account that only will need to be authenticated once. This process is done by a Staff Member.

4.3 Start Menu

When the application starts, the start menu appears. As it can be seen in Figure 4.1, none of those options are oriented to the Regular Users. Only Staff Members will interact with this menu. By first registering and authenticating themselves, these users can register and authenticate children, which are the Regular Users of the application. The hierarchical navigation of this screen has only one level, so each button takes the user to only one respective page.

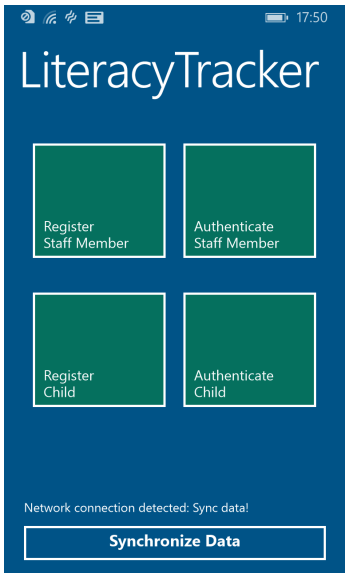


Figure 4.1: Start menu screen.

To register or authenticate a child, a Staff Member must be authenticated first, or a message dialogue will appear. The registration and authentication pages are robust enough to avoid and alert users to some common validation errors, such as leaving empty fields.

When authenticating children, the staff member has one field to specify if the child has ever been tested in the traditional pen-and-paper test. That option was implemented to correlate the results with the LiteracyTracker assessment results.

Due to the internet connectivity barriers in the SSA, the application will advise the Staff Member when the device is connected. If that is the case, the "Synchronize Data" button will appear together with a message. If not, they both will stay hidden. In the first case, the user taps the button and the data will automatically be uploaded to the cloud account. In the first synchronization, the user will need to authenticate in the cloud account with the email account and respective password. After that moment, the process will no longer be needed since the application will save the credentials. Thus, only one tap is needed and after a few seconds a message will appear on the screen saying the data was already uploaded, as shown in Figure 4.2.

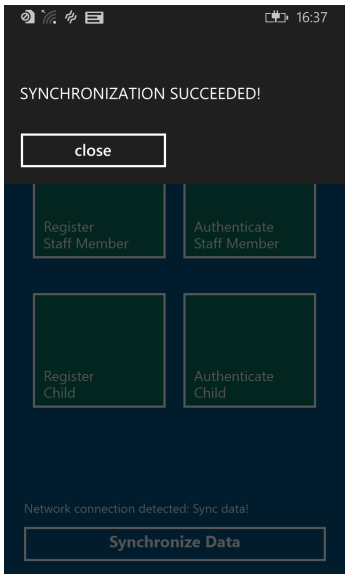


Figure 4.2: Message of succeeded data synchronization to the cloud storage.

4.4 Children's Home Screen

After the authentication process conducted by the Staff Member, the application shows the children's home screen. Starting from this moment the active user is now the child. As it is showed in Figure 4.3, these users have the first contact with the avatar, which appears together with a salutation message with the name of the user. They have three options: take a new assessment test, see their rewards list and exiting from their session.

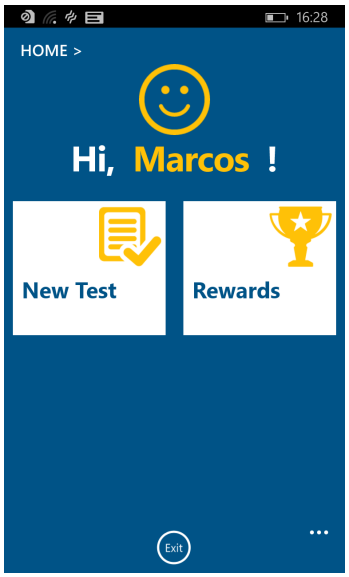


Figure 4.3: Children's home screen.

In Appendix B, screen of the registration and authentication process are shown.

4.5 Literacy Test

This section presents all the design related to the literacy assessment module. For the implemented prototype, only five of the tasks of the proposed literacy model, showed in Table 3.2, were used and therefore represented as levels. Their name format follows the sequence order by which they appear. Table 4.1 shows the ordered list of levels, as well as the corresponding tasks.

Table 4.1: The ordered list of implemented levels.

Name	**Task**
Level 1	Writing the alphabet letters
Level 2	Word Recognition
Level 3	Vocabulary Knowledge
Level 4	Writing words from dictation
Level 5	Comprehension of sentences

With this set of five exercises, reading and writing skills were granted, along with the listening comprehension. Oral reading fluency was not implemented. As mentioned in [SPM+09], the speech recognition system like the one that would be required to implement that task, inevitably has to be highly accurate to reach the objectives of the application. It would be too expensive, relying on the time needed, to ensure that this condition would be satisfied in LiteracyTracker, and for that reason this was left for future work.

For each level, there are a certain number of rounds of questions and in each round a new question appears with a new set of possible answers. When the rounds are over, if the users still have "lives left", then they will level up. Otherwise, as soon as the "lives" are over, the test is terminated.

By the end of each test, another JSON file is being generated by the application with all the data gathered during the assessment, such as the correct and incorrect given answers, the time spent in each level, points earned, among others. Its access is deliberately hidden from all the users, since it is saved in the internal storage of the device. Then, when the application data is synchronized to the cloud account, these files can be analysed.

Each level has an XML database file with configurations and associated exercise contents.

In Appendix A, two examples of the mentioned JSON and XML files are showed.

4.5.1 Status Header Panel

The current status of the literacy assessment is shown in every level screen, through an informative header panel, such as the one represented in Figure 4.4 in which the following elements are always present:

- the navigation path (starting from the user's home screen) ;
- the name of the level;
- the first name of the user;
- the user's lives left (the number of wrong answers that the user can give in each level);
- the first name of the user;
- the avatar;

Figure 4.4: Example of the status header panel.

Regarding the avatar, the six possible states of its mood were implemented. Figure 4.5 shows those states, as mentioned before in Section 3.4, by this order : neutral, happy, sad, awesome, awful and concentrated. The last one is only used when the adaptive mode interface is enabled, as Subsection 4.7 refers. The hand-drawn like style of the character was a requirement met, since research proved to be better understood by illiterates than other kind of visual representations styles [MPT05, MPT07].

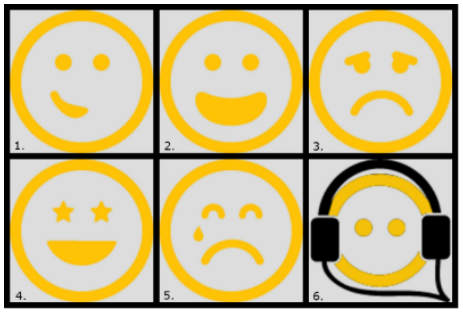


Figure 4.5: The six states of avatar's mood during the literacy assessment.

4.5.2 Command Bar

The command bar always appears on the bottom of the screen. It is a customizable native bar from the Windows Phone devices. During the assessment, it contains always two icons: the home icon and the help button. They both are labelled with a description, however the only way to see that is by tapping the "three dots" icon, also integrated in the command bar. Figure 4.6 shows the two icons with the respective label descriptions. Tapping the first one cancels the assessment and the user is taken to the home screen. The second one provides audio help about the level. Usually, a

description of the level is heard. However, when the adaptive interface mode is enabled, the button will repeat an explanation of the level, step by step.

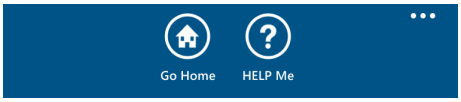


Figure 4.6: When the "three dots" button is tapped, the textual description of the icon appears.

4.5.3 Feedback

The visual feedback for correct or incorrect answers uses two colours: green in case of success, or red otherwise. The visual feedback suffered some changes and refinements during the design phase. Coherence between levels was another desired requirement to meet, regardless of the type of action that the level would require from the users. For that reason, the green and red colours were only used for visual feedback situations. These situations can be seen in the next sections, where each level is detailed.

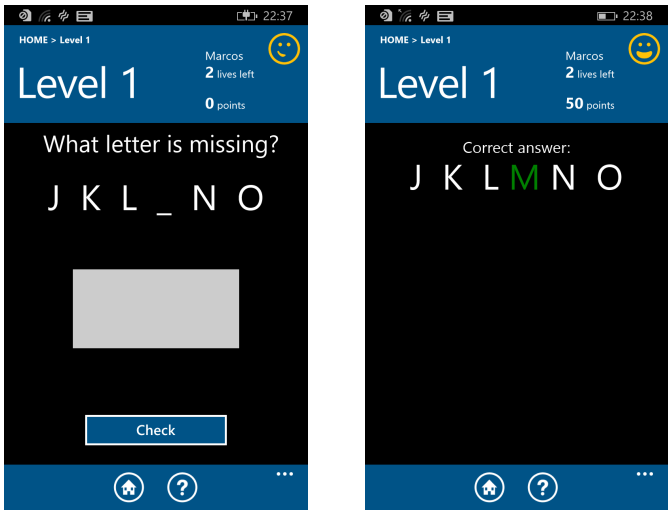
Also, as mentioned before, the avatar element changes its state as well.

4.5.4 Level 1

The first level of the literacy assessment asks the user for a missing letter of one broken sequence of the alphabet. As Figure 4.7a shows, the missing letter is the one who should appear instead of the underscore symbol. To answer the question, the users should write in the answers box and then validate the given answer by touching the "Check" button. To prevent errors, the application disallows the writing of words and accepts both capital and small letters as a possible solution. If the answer is correct, the full sequence of letters is shown with the missing letter coloured in green, as Figure 4.7b shows. Otherwise, the same happens but with the letter coloured in red.

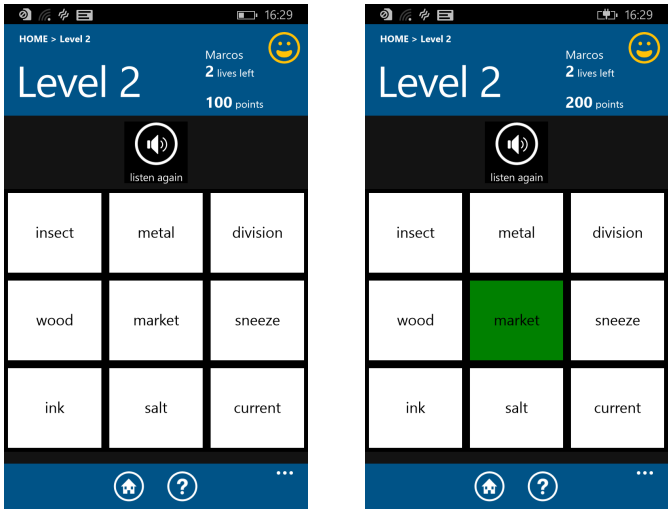
4.5.5 Level 2

As soon as the *Level 2* starts, a list of words is displayed. At the same time, the sound of one of those listed words is heard. This level database contains a list of 400 general words of the basic English language [Ogd44]. Every time the users want to listen the word again, they can tap the "listen again" button visible on the top of the grid of words. Each grid will turn green or red as the answers are correct or not. Figure 4.8b shows an example of the first case. At the same time the colour is displayed, the sound of the selected word is played. Thus, if the answer was incorrect, the user will assimilate the selected word and notice the differences.



(a) Starting screen of *Level 1* (b) Correct answer in *Level 1*.

Figure 4.7: Examples of screens of *Level 1*.

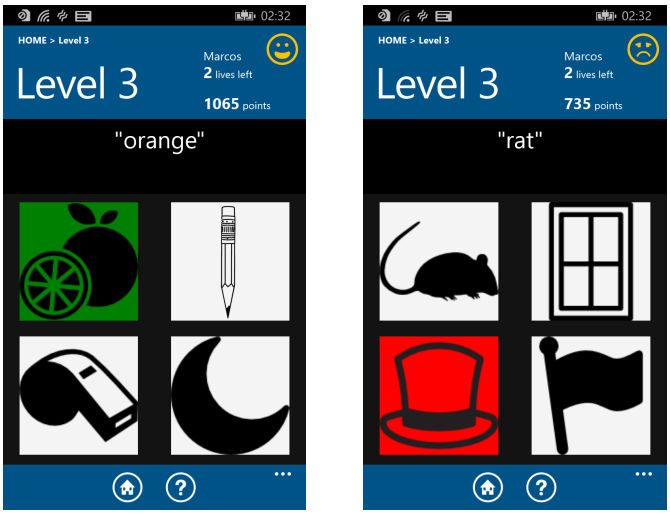


(a) Starting screen of *Level 2* (b) Correct answer in *Level 2*.

Figure 4.8: Examples of screens of *Level 2*.

4.5.6 Level 3

The objective of *Level 3* is to evaluate if the users know the basic meaning of words. To assess that, the level displays one word and a set of four images. The users have to match the word with its pictorial representation. According to [Ogd44], there is also a list of 200 *picturable* words, i.e., vocabulary words for things that are capable of being represented by a picture. In this level, only 119 of those words were used, mainly because the correspondent images had to be evident for the users and, at the same time, the images had to be license free, which difficult the process of choice. Figure 4.9 shows the visual feedback for both correct and incorrect answers.

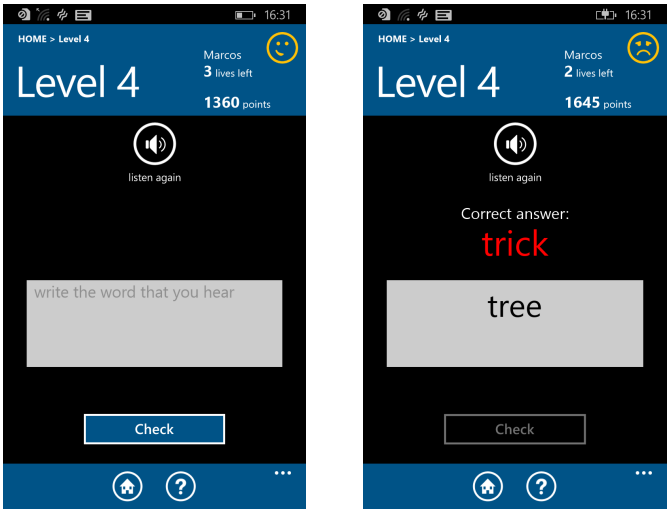


(a) Correct answer in *Level 3* (b) Incorrect answer in *Level 3*.

Figure 4.9: Examples of screens of *Level 3*.

4.5.7 Level 4

Regarding *Level 4*, the dynamic of that exercise is similar to the *Level 2*, however this time the users have to write the listened word and then validate the answer by tapping the button "Check". The answers box has the description of the required action. In that way, the users can focus on the "listen again" button and the screen is kept clean. If the given answer was incorrect, the application shows the solution coloured in red, as it is possible to observe in Figure 4.10b. Otherwise, as expected, it will be green. The database of words used in this exercise was the same as the used in *Level 2*.

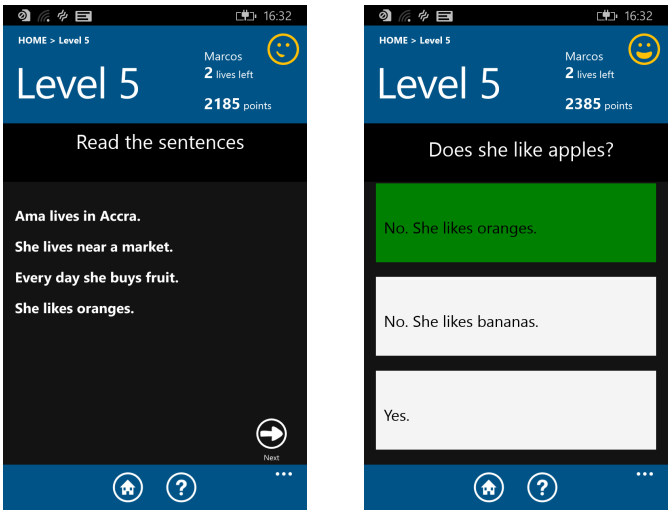


(a) Correct answer in *Level 4* (b) Incorrect answer in *Level 4*.

Figure 4.10: Examples of screens of *Level 4*.

4.5.8 Level 5

The last level implemented is divided into two parts. First, as shown in Figure 4.11a, the users have to read four short sentences that are related between themselves. Then, after tapping "Next", a question related to the read sentences is displayed as well as three possible responses. Again, the visual feedback based on green and red colours are displayed after the tapped answer. Regarding the database, this time it was fulfilled with exercises gathered and adapted from scholar books.



(a) Sentences

(b) Question and possible answers

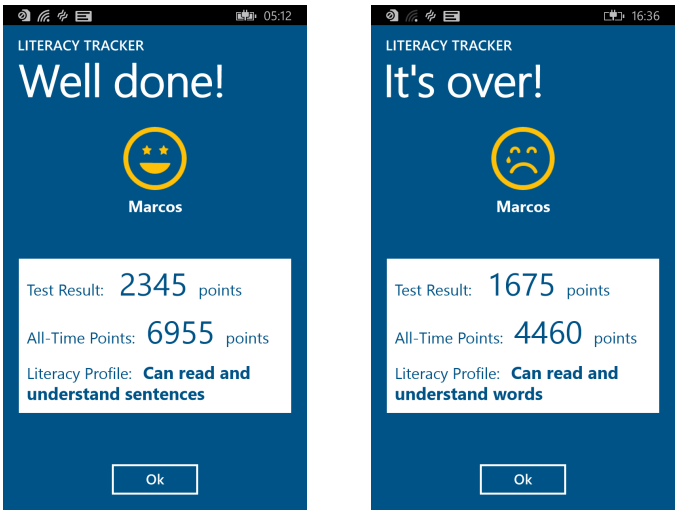
Figure 4.11: Example of screens of *Level 5*, showing the sentences, the question and possible answers.

4.5.9 Final Results

By the end of each test, as Figure 4.12 shows, two possible screens may appear regarding the user's performance. When the users complete all of the assessment, a screen is shown with the avatar in happy mood. In case they terminate it but do not complete it, another screen appears but with the awful mood. Whatever the case is, both will appear with the test result, the all-time points earned and the literacy profile relative to that test. The all-time best score, whose value would be used to rank users in the individual leaderboard, was not implemented, since the leaderboards were not implemented either.

4.6 Rewards

Each Regular user has a list of rewards to unlock, when some challenges are achieved. As mentioned in Subsection 3.5.1, it was specified that they should hold symbolism and meaning to the target users. There are over 100 ethnic groups living in Ghana [Kwi15]. For that purpose, a research was made in order to find possible elements of the Ghanaian culture recognized by all the



(a) Congratulatory screen. (b) Defeat screen.

Figure 4.12: Two examples of screens that may be displayed when the assessment is over.

population, to use as rewards. Some African symbols known as *Adinkra* which are ubiquitous in Ghana, were discovered [Qui15, Tet06, Mac15, Kou15].

Adinkra symbols were originally created by The *Akan* of Ghana and the *Gyaman* of Ivory Coast, both ethnic groups in West Africa. They have a decorative function, being extensively used in fabrics, pottery, logos and advertising, but also used to communicate evocative messages that convey traditional wisdom, aspects of life or the environment, rules and life in society [Tet06, Kou15].

There are 63 of these symbols which are well catalogued. However, for this prototype only 10 were selected for the purposed challenges. However, following the literature review, specially the guidelines proposed in Section 2.3.1, symbols whose meanings could be related with gender distinctions, political differences or religious beliefs, for example, were avoided.

Table 4.2 lists the chosen *Adinkra* symbols, their meanings and associated challenges that users have to achieve in order to unlock them.

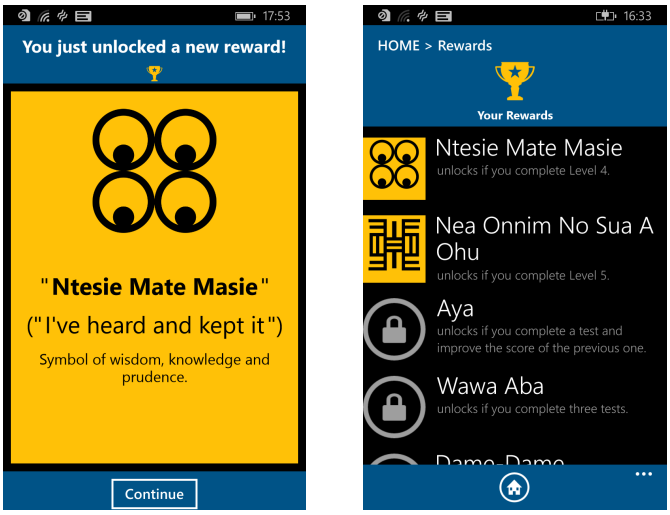
In [ASH13], it is explained that providing detail to learners about how to earn the badges can mitigate some negative motivational effects while preserving the assessment goal of badges. For that reason, as shown in Figure 4.13, for each reward in the list, there is a label detailing the challenge that is need to achieve in order to unlock each one of the badges.

In Appendix B is shown more examples of unlocked rewards.

Every time the users unlock a new reward, a new screen is displayed to them, showing what they just accomplished, as shown in Figure 4.13a. Purposely, to avoid confusions and misconceptions from the users about gamification elements whose names could appear on the screen, such as challenges, achievements or badges, the only thing that it is transmitted to the user is the word "Reward". That way, it is expected that it is easier for the users to understand that by succeeding in a certain skill or set of skills, they would be rewarded.

Table 4.2: List of *Adinkra* symbols used as rewards and respective meanings and associated challenges.

Symbol	Name	Symbolism	Translation	Challenge
	*Nkyinkyim*	initiative, dynamism and versatility	Twisting	complete *Level 1*
	*Akofena*	courage, valor, and heroism	Royal sword	complete *Level 2*
	*Woforo Dua Pa A*	when you work for a good cause, you will get support	"When you climb a good tree"	complete *Level 3*
	*Ntesie Mate Masie*	wisdom, knowledge and prudence	"I've heard and kept it"	complete *Level 4*
	*Nea Onnim No Sua A Ohu*	knowledge, lifelong education and continued quest for knowledge	"He who does not know can know from learning"	complete *Level 5*
	*Aya*	endurance and resourcefulness	Fern	complete a test and improve the score of the previous one
	*Wawa Aba*	skillfulness, hardiness, toughness and perseverance	Seed of the wawa tree	complete three tests
	*Sankofa*	importance of learning from the past	Return and get it	complete a test without making mistakes
	*Nyansapo*	wisdom, ingenuity, intelligence and patience	Wisdom knot	complete three tests without making mistakes
	*Adinkrahene*	greatness, charisma and leadership	Adinkra king	earn all the rewards



(a) Example of unlocked reward.

(b) List of rewards, some of them still locked.

Figure 4.13: Example screens of a reward and the rewards list.

4.7 Adaptive Interface

The adaptive interface mode was implemented, however only some of the necessary specifications to prove the concept were completed. When the user finishes one literacy assessment, the application will calculate if he should be assisted by an adaptive interface or not, based on the criteria specified below.

1. **Behaviour analysis:** whatever the previous interface mode was, the application is permanently counting the number of times the users resorted to the help button, placed on the command bar, and in which level they tapped it. Adding to this, in each level, the amount of time that users took to answer and to level up is also counted. These factors can be indicators of possible incomprehensions and difficulties of the users, while they interact with the system.

2. **Literacy profile:** the assessment performance is also used for the calculation.

3. **Final Formulation** the suggestion will appear if one of these situations is valid:

   - users couldn't reach *Level 3*;
   - users spent more than 25 seconds in *Level 1* or 1 minute in *Level 2*;
   - users tapped twice, or more, in the help button;

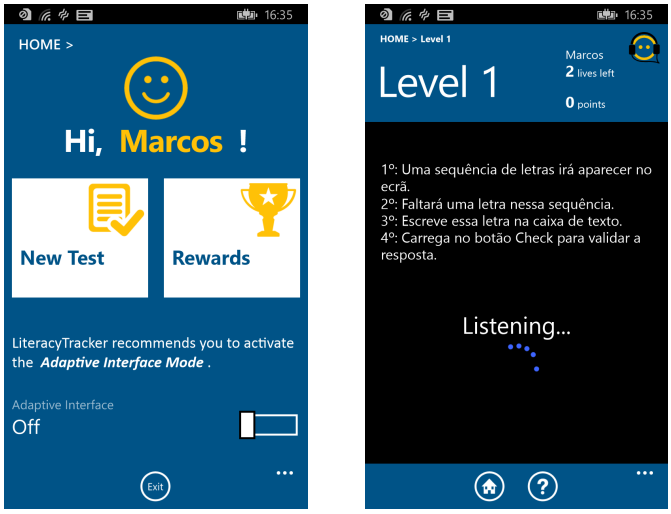
Regarding the standard values, they were used in the evaluation processes of the prototype. However, as mentioned before, they could be changed manually on the XML configurations file.

Figure 4.14a shows what happens when the calculation result is the adaptive mode. Contrary to the specified system, the prototype implemented shows the recommendation on the Regular

Users home page, which implies that the Staff Member has to interact on that screen too. From the moment a user starts an assessment, until the end of it, the navigation structure of the system is linear. In that way, after the end of the assessment, the application goes back to the home screen of the user and the recommendation will appear. Nevertheless, the switch button is set off by default, since it is still expected that the staff member will take the decision.

Another feature of the adaptive mode is presented in Figure 4.14. As it is possible to observe, the users can have a step by step textual explanation of the level in their native language. At the same time, as the avatar indicates, the user is also listening to the instructions. Along with the avatar, a textual reference and a progress ring appear on the screen while the audio is transmitted, warning the users to the fact that the application is "speaking" to them.

Adding to this, the question of *Level 1*, which can be seen in Figure 4.7a, is removed so as to make the interface cleaner and the size of "listen again" button of *Level 2* is increased.



(a) Adaptive interface mode recommendation.

(b) *Level 1* being explained befor it starts.

Figure 4.14: Examples of screens about the adaptive interface mode.

4.8 Summary

This chapter showed the implementation of the proposed solution, with its design and technological decisions. In Section 4.1, some initial considerations were given to better understand the remain sections. Section 4.2 listed all the system requirements. From the Section 4.3 to the Section 4.7 all the implementation decisions are explained, namely all the levels design process and the chosen reward symbols.

Chapter 5

Evaluation and Results

This chapter details both evaluation methodology and results. By analysing these results it is possible to evaluate the implemented system, as well as to identify the failures that need to be corrected in order to obtain a more accurate solution. The evaluation sessions that were conducted, the participants and the procedures of each test are detailed in the next sections.

5.1 Evaluation Overview

The main purposes of the evaluation process of the system were five:

1. to refine the design and usability based on the evaluation observations;
2. to validate the designed literacy test model and its database of exercises;
3. to check if gamifying literacy would bring the desired motivation and fun to the participants;
4. to figure out the effect and the feasibility of the adaptive interface;
5. to assess participants with the basic VPWA's literacy test and then with the LiteracyTracker and compare the results;

As mentioned before in Chapter 4, the literacy test model was created to assess the English language. For that reason, one interview with an English teacher of the one of the schools was requested in order to validate the content of the literacy test. In that interview, the importance of the designed literacy model was explained along with the project's context. Also, a demonstration of the solution was presented. In the end, the teacher praised the simplicity of the exercises format, and highlighted the matching levels of words with images and sounds. Moreover, she gave suggestions for possible more levels.

The tests were conducted in three different sessions, in Portugal. The participants were students aged between 7 and 12, from three Portuguese schools, places where the tests were conducted. The fact that English was not their first spoken language allowed to simulate illiterate and semi-literates users, since they had different levels of knowledge of the English language.

Based on their results, one improved version of the LiteracyTracker should be achieved to be later tested on Ghana, due to a partnership between Fraunhofer Portugal AICOS and the VPWA organization.

5.2 Evaluated Users

In total, 24 individuals participated in the evaluation process. As it is shown in Table 5.1, 13 of them are regular smartphone users and none of them have Windows Phone devices. In fact, all the 13 revealed to have Android smartphones. In this respect it is worth mentioning again that the developed solution runs Windows Phone devices.

It is also possible to observe that the participants are all children within the age range of 7 to 12 years.

Table 5.1: General information about all the tested participants.

Participants information	
Number of participants	24
Age range	[7 - 12]
Number of smartphone users	13
Windows Phone users	0

The gathered sample of participants contains children that enrol both first and second cycles of basic education. However, by the time the tests were conducted, the Portuguese educational system stipulated that the English language was only mandatory from the second cycle [San15]. This means that the sample can be divided into two groups of study, to better analyse the levels of knowledge of the English language by the participants.

Table 5.2 and Table 5.3 display the sample respectively divided into two groups: mandatory and optional English learners.

Table 5.2: Representation of the children attending the 2nd cycle, whose English language is integral part of their course syllabus.

Group 1: mandatory English learners	
Number of participants	8
Age range	[11 - 12]
School year	6th
Number of years learning English (mandatory)	2
Number of smartphone users	8
Number of Windows Phone users	0

As Table 5.2 shows, only 8 of the 24 participants, aged between 11 and 12 years, attend mandatory English classes and all of them for two years, since they are in the 6th year of school.

Below, in Table 5.3, the second group of the sample is represented by the participants who are optionally learning English. These children are aged from 7 to 10, attending the 1st cycle of basic education.

Table 5.3: Representation of the children attending the 1st cycle and learning the basics of English as their own will.

Group 2: optional English learners	
Number of participants	16
Age range	[7 - 10]
School years	[1st - 4th]
Number of years learning English (optional)	[1 - 4]
Number of smartphone users	5
Number of Windows Phone users	0

To better understand the distribution of these participants by the years of English language learning, Table 5.4 should be consulted.

Table 5.4: Distribution of participants, of the Group 2, by the years of English language learning.

Years of learning / **Age**	**1**	**2**	**3**	**4**	**Participants by age**
7	*3*	*1*	*0*	*0*	4
8	*0*	*2*	*2*	*0*	4
9	*0*	*0*	*2*	*0*	2
10	*0*	*0*	*2*	*4*	6
Participants by years of learning	3	3	6	4	16/16

5.3 Test Procedures

This section details the taken test procedure. To better understand it, the section is divided into three moments: before, during and after the test.

5.3.1 Test Preparations

The three test sessions took place in different schools at distinct days. For each case, it was ensured that the test session would be conducted in a quiet room, isolated from any external perturbations. Also, parental authorization was requested to the children in order to be able to participate.

The equipment required for each session is listed below:

- 4 Windows Phone 8.1 devices:
  - Nokia Lumia 1020, with 4.5 inches of display size [Gsm15a];

  - Nokia Lumia 830, with 5 inches of display size [Gsm15c];
  - Nokia Lumia 630, with 4.5 inches of display size [Gsm15b];
  - Nokia Lumia 920, with 4.5 inches of display size [Gsm15d];

- 4 earphone kits;
- 1 digital video camera;

Each smartphone was placed together with one kit of earphones. The earphones were used in order to ensure the participants would listen the audio instructions The video camera was positioned in a way that would be possible to record the user's hands interacting with one of the devices, since only one camera was available.

Next, before the test starts, a brief and simple explanation of the main objective of the project was presented to the participants. Due to the age of the participants, it was made clear that the goal was not to assess how clever they could be and that the results would not be taken in account for school evaluation purposes.

Finally, each user was registered and authenticated in the application and a description of the task they would have to do was given.

5.3.2 During the Test

Due to time restrictions, the first school established that four participants at a time should be tested. For that reason, in the first session it was not possible to assess the literacy levels of the participants with the literacy test model of VPWA, as planned. In the other two sessions, however, that was not the case.

Only one task was introduced and asked to the participants: to start a new literacy assessment and keep answering until the test terminates.

During the test, notes were taken and videos were recorded for further analysis.

5.3.3 After the Test

After each participant was tested, a System Usability Scale (SUS) questionnaire followed by an additional survey composed by only three questions about the gamification elements present in the system was given to be answered. SUS is a simple, ten-item *Likert scale* giving a global view of subjective assessments of usability [Bro96]. Each item has five response options for respondents: from "1 - strongly disagree" to "5 - strongly agree".

Interpreting the scores of SUS can be complex and despite the wide usage of SUS, there has been little guidance on interpreting SUS scores [Usa15b, Sau15]. In [Sau15], the scoring process is explained as below:

- for odd items: subtract one from the user response;
- for even-numbered items: subtract the user responses from 5;

- this scales all values from 0 to 4 (with four being the most positive response);
- add up the converted responses for each user and multiply that total by 2.5. This converts the range of possible values from 0 to 100;

Although the scores are ranged from 0 to 100, they should not be interpreted as percentages. Based on research, a SUS score above a 68 would be considered above average otherwise is below average. Often they are normalized to produce a percentile ranking [Usa15b, Sau15].

However, in [BKM09] a study was done in order to help to determine what individual SUS scores mean by adding an adjective rating scale. Figure 5.1, shows a comparison of the adjective ratings, as well as other common used scales, in relation to the SUS scores, which can help to better understand the results.

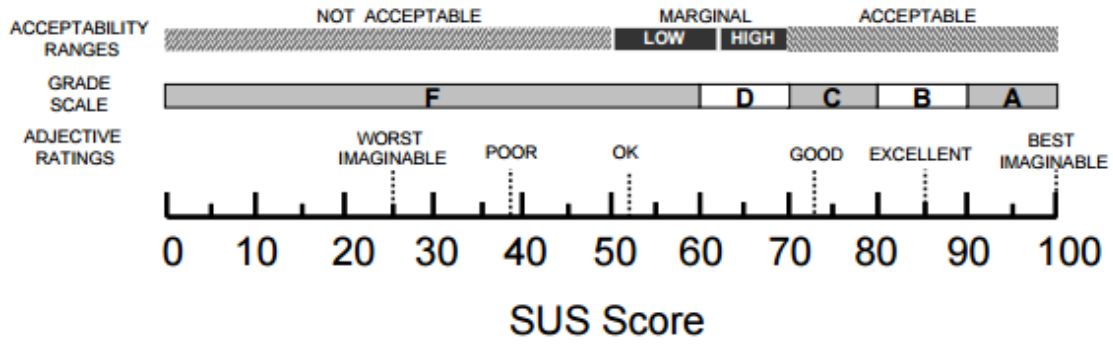


Figure 5.1: A comparison of the adjective ratings, acceptability scores, and school grading scales, in relation to the average SUS score [BKM09].

After the first session, it was concluded that the SUS was too verbose and hard to understand for some of the participants. Thus, as in the next sessions some of the participant still were learning the basics of Portuguese language, the SUS was adapted with a more simplified language but without compromising the main objectives of the questions.

The second small survey was composed by only three questions, which are translated and listed below:

1. *Q1*: I found the different types of exercises presented were interesting.
2. *Q2*: I found the presence of points and lives in each level interesting.
3. *Q3*: Reviews, difficulties and suggestions (optional).

In Appendix C, the SUS questionnaire is shown as well as the second survey used, as well as all the calculated scores for each session.

The first two were classified with the same scale of SUS. The last one, was an optional open answer question.

5.4 Evaluation Sessions

Since the developed solution followed an iterative process of design, when problems were found in user testing sessions, they needed to be fixed. This cycle of design, test and redesign, was repeated for the three sessions expecting to meet the usability goals. For each session, video recordings, annotated observations, questionnaires and feedback from the users were taken into consideration for the refinement of the system. Also, the data of the users performance during the period of interaction, internally gathered by the application, was analysed with the same purposes.

5.4.1 Session 1

The first session of user testing occurred for about one month after the implementation process started. All the elements of the Group 1, mentioned in Table 5.2, participated in this session. The main purpose of this test session was to test the usability as well as the navigation and information structure of the prototype.

As requested by the school, due to time restrictions, they were divided into two groups of four and in each group the individuals were tested at the same time. Three captures of video were taken.

Regarding gamification elements, only levels, points and the avatar were implemented. By that time, the prototype already had implemented the five levels of the literacy assessment model. The order of the levels, however, was different: *Level 4* and *Level 5* were exchanged at the time.

Moreover, achievements or rewards were not tested. Nevertheless, an opportunity emerged to observe how the motivation and engagement levels of the participants would be during the tests, since no reward system was implemented.

5.4.1.1 Detected Anomalies

As it was expected, the first of the sessions exposed more defects. The list presented below shows the meaningful ones:

- ***Level 1* was not well understood:** in terms of difficulty level, *Level 1* was clearly the less exigent one. As Figure 5.2 shows, it was required the missing letter by the description of the exercise. Even a the answer box has a placeholder text, indicating that the user should type there the answer. However, many users took more than 15 seconds to understand and answer for the first time.

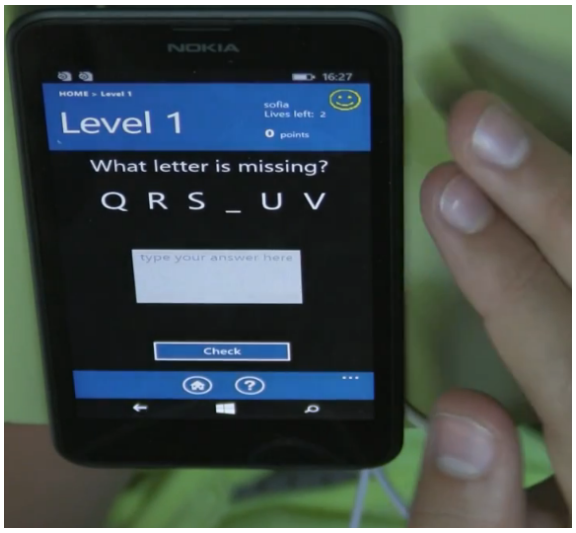


Figure 5.2: In Session 1, *Level 1* took a lot of time to be understood by some users.

- **Empty answers:** one user accidentally tapped the "Check" button before writing the answer, in *Level 1*. The application have wrongly assumed the empty answer as a wrong answer, deducting points to the user.

- **Enabled tapping while displaying visual feedback:** one user quickly tried to correct one wrong answer by tapping another one, while the visual feedback was still displaying the red color for the first answer. The application wrongly returned double feedback for both answers. In Figure 5.3, a similar case happened when, in *Level 3* one user tapped the screen with the anxiety to move to the next question. The visual feedback of the first answer blocked until the end of that level.

Figure 5.3: In Session 1, the tapping action was enabled while the application was displaying visual feedback.

- **Lack of information in *Level 4*:** In the, at the time, *Level 4*, after the users read the sentences it was expected that they would hit the *Next* button to move toward the second part of the exercise. Some felt lost for some seconds until realize the "Next" button was in the bottom command bar. Even when performing the level for the second and third times, some users tapped first in the middle of the screen before quickly tap the right button, as shown in Figure 5.4. However, one case was even worst: one user felt so confused that exclaimed in loud voice that the test was over.

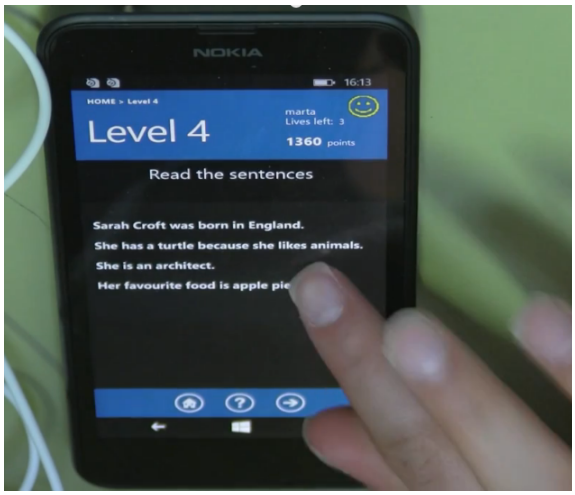


Figure 5.4: In Session 1, the user assumed that tapping somewhere on the screen before realizes the action is done by an a button.

5.4.1.2 Additional Observations

In spite of the fact that no reward system was presented to the users, motivation and engagement levels of the users were observable. In fact, it was agreed with the school that only one test per student would be done. However, in more than one case, once the participants terminated the first attempt immediately asked if it was possible to repeat the test in order to improve their score. One user, after beating on the table to release the frustration felt for the final test result, just started a new one without asking any permission. After seeing the scenario, the teacher, which was also present in the room, allowed the participants to perform other tests. Only 1 of the 8 participants didn't perform more than one test. That was a good indicator of the enjoyment the participants were feeling.

5.4.1.3 Results

The results of the SUS questionnaires for Session 1 are presented in Table 5.5. It is observable that the mean value, 91.56, which means that the usability of the prototype of Session 1 can be classified between "Excellent" and "Best Imaginable", according to Figure 5.1.

Table 5.5: Global SUS score of the Session 1.

Session 1	
User	SUS Global Score
U001	97,5
U002	82,5
U003	82,5
U004	95
U005	90
U006	92,5
U007	97,5
U008	95
Mean	91,56
Standard Deviation	6,11
Minimum	82,5
Maximum	97,5

Regarding the second questionnaire, Figure 5.5 shows the results of each user for the questions *Q1* and *Q2*. The range scale used was from "1- Strongly disagree" to "5-Strongly agree". Even without the presence of rewards, in general users appreciated the implemented gamification elements, such as points and levels.

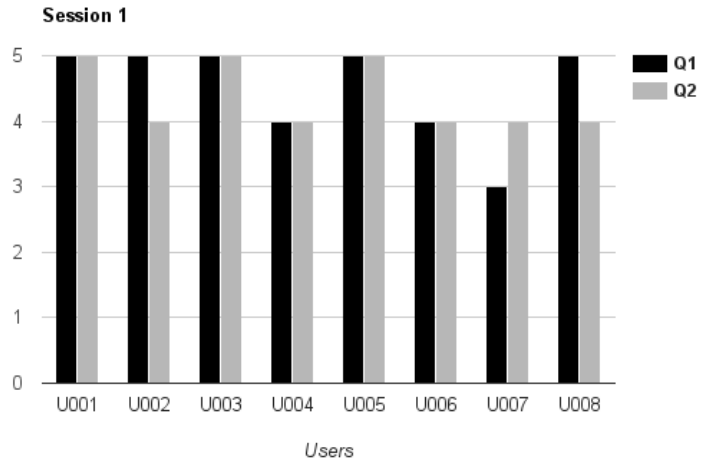


Figure 5.5: Results of the second survey of the Session 1.

5.4.2 Session 2

From the Group 2, mentioned in Table 5.3, 9 students participated in this session, that occurred for about one week after the first one. This time, the alignment of the levels was already correct. The adaptive interface mode was already implemented and it was tested with the textual explanation

of the levels in the native language of the participants. However, the audio explanation was still in English. Regarding the detected anomalies in Session 1, they were fixed. However, in *Level 1* only the text inside the answers box was removed, just appearing when users tried to validate empty answers. Also, the size of the answers box was reduced to give the impression that just a letter would be required in that exercise. Even knowing the participants of Session 1 revealed some difficulties to understand the level, the layout of this level didn't change too much purposely. The aim was to see if the adaptive interface mode would help the users.

Regarding the open answer question of the second survey, one user suggested the presence of more pictorial representations in the application, aside from the existence of the avatar.

5.4.2.1 Detected Anomalies

- ***Level 1***: the *Level 1* was a problem again to be understood by the younger participants, regardless its low level of difficulty, during the first attempt. The older participants, however, didn't face the same problem.

- **Short duration time of the explanation**: the adaptive interface mode was well perceived, since that in the moment it was enabled the participants understood something was different at the beginning of the *Level 1*. The problem was the duration time of the explanation was too short for the younger participants: they couldn't read so fast.

- **The explanation of the levels couldn't be stopped:** as it is possible to observe in Figure 5.6, in the adaptive interface mode, at a certain moment, one of the users tried to stop the explanation of the level by tapping the progression ring. This user already had assimilated what it was supposed to do and wanted to advance to the exercise but the application wasn't ready for it.

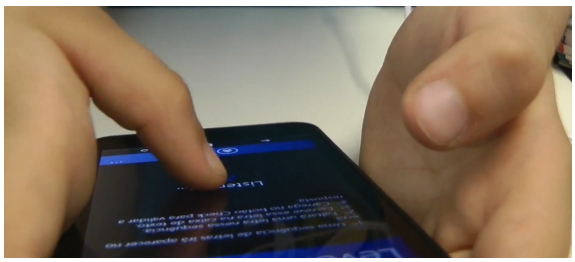


Figure 5.6: User trying to stop the explanation of the level by tapping the screen(adaptive mode interface), in Session 2.

- **Header panel information**: some users showed surprise when they lost, like they didn't understand or saw the information about the number of "lives left" they had at each level, unlike the points earned, for example. As it is possible to see in Figure 5.7, the size of the number of lives was too small and less emphasized, comparing to the size of the number of points.

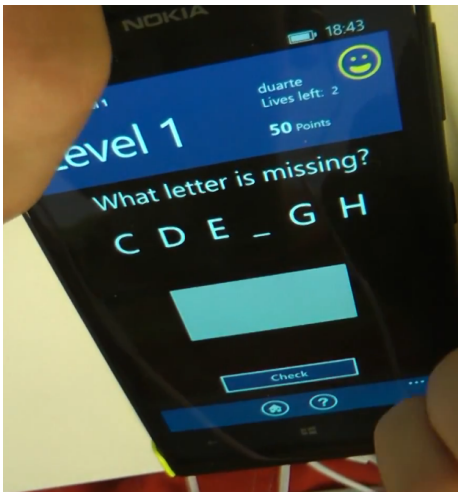


Figure 5.7: In the implemented prototype of the session 2, the information of lives of the user was not so emphasized as it should, in Session 2.

5.4.2.2 Additional Observations

In this session was possible to observe a healthy competition between two of the participants, that was recorded in the videos. While doing their tests, they often communicate between each other to ask in what level they were as well as the number of points they had.

Moreover, as the video observations and the open answer questions of the second questionnaire revealed, the younger users felt difficulty in understanding the *Level 1*, however the following levels were perfectly understood. For the younger participants, the duration of the explanation should be controlled by them. Also, the audio explanation could be provided in their mother tongue if possible.

Regarding the open answer question of the second survey, some users asked for more different "games", which was a reference to the levels of the assessment.

5.4.2.3 Results

The results of the SUS questionnaires for the Session 2 are presented in Table 5.6. This time, and using again the same adjective rating scale, usability of the prototype is classified between the "Good" and the "Excellent", since the mean SUS score of the Session 2 was 75. One possible explanation for this decreasing value is the fact that, in Session 2, almost 45% (4 out of 9) of the participants were enrolling the first and second years of school. In contrast, in Session 1 all the participants were from the sixth year of school. That can explain why even with the adaptive mode interface implemented the users had some difficulties to understand some levels of the application. Even the explanation of the levels in their native language was not a good help, since they weren't able to read Portuguese fluently. In fact, it was visible the frustration of some of the them when they weren't able to read the instructions as fast as the explanations was given.

Table 5.6: Global SUS score of the Session 2.

Session 2	
User	SUS Global Score
U009	67,5
U010	80
U011	80
U012	80
U013	80
U014	97,5
U015	77,5
U016	65
U017	47,5
Mean	75,00
Standard Deviation	13,81
Minimum	47,5
Maximum	97,5

Regarding the results of the questions *Q1* and *Q2* of the second survey, Figure 5.8 shows that, besides the user *U012*, every participants strongly agreed with the different type of exercises as well as it found interest on the presence of points and lives in each level of the application.

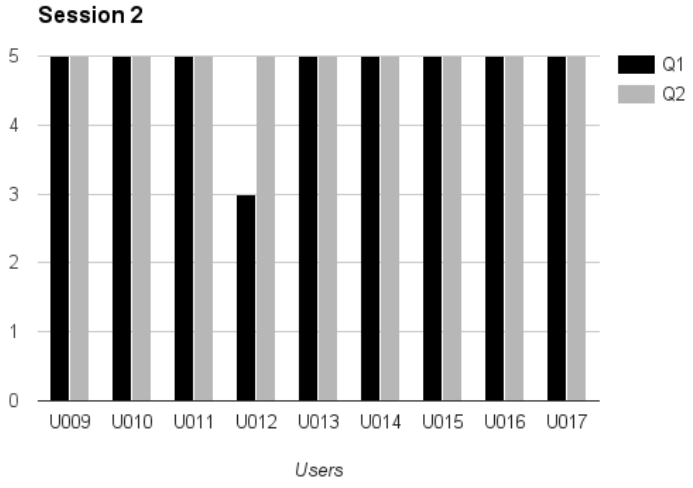


Figure 5.8: Results of the second survey of the Session 2.

In Table 5.7, it is possible to compare the results of each participant when assessed by the traditional test of VPWA or by the LiteracyTracker. It is important to notice that the identification of each user in this table is distinct from the questionnaires, since they were answered anonymously. It should also be noted that the exercises "word recognition" and "comprehension of sentences" of

the basic literacy test are similar to *Level 2* and *Level 5* of LiteracyTracker. Also, the basic literacy test is no more than a pen-and-paper based test, so the exercises are static, i.e., the user will answer for four or five times based on the same set of words that he is seeing, while in LiteracyTracker the exercises are dynamic. The set of words changes for each new question, as well as the possible answers. This situation avoids any kind of memorization or elimination processes by the users.

Table 5.7: Comparison between the basic literacy test of VPWA and LiteracyTracker, with results of the participants of Session 2.

	SESSION 2				
	Basic Literacy Test			**LiteracyTracker**	
	Exercise completion		**Literacy Profile**	**Level Reached**	**Literacy Profile**
ID	**Words Recognition**	**Comprehension of sentences**			
ID201	yes	yes	**can read and understand sentences**	*Level 4*	**can read and understand words**
ID202	yes	no	**can read words**	*Level 2*	**can read and write alphabet letters**
ID203	yes	no	**can read words**	*Level 2*	**can read and write alphabet letters**
ID204	yes	no	**can read words**	*Level 3*	**can read words**
ID205	yes	no	**can read words**	*Level 4*	**can read and understand words**
ID206	yes	no	**can read words**	*Level 3*	**can read words**
ID207	yes	no	**can read words**	*Level 3*	**can read words**
ID208	yes	no	**can read words**	*Level 2*	**can read and write alphabet letters**
ID209	yes	no	**can read words**	*Level 4*	**can read and understand words**

The literacy test of the VPWA assesses users with the word recognition exercise and if they succeed they are categorized as "can read". Next, they are evaluated with the exercise about comprehension of sentences and if succeed, the users are categorized as "can read and understand sentences". Regarding the LiteracyTracker, none have succeeded in *Level 4* (writing the listened word). Two of them (*ID205* and *ID209*) proved that can read and understand the meaning of the words, something that would be impossible to say with the basic literacy test, since there are no exercise oriented to evaluate that. Only three of the participants (*ID204*, *ID206* and *ID207*) were categorized with the same profile in both tests

Finally, by analysing the JSON files of each assessment, it was possible to conclude that only 3 users used the help button of the command bar. In *Level 1* the button was tapped once by *ID207* and twice by *ID205*. In *Level 2*, *ID206* tapped the help button for 4 times. Each one of them terminated their assessment in a higher level, as Table 5.7 indicates.

5.4.3 Session 3

The remaining 7 individuals from the Group 2 participated in the Session 3. The time difference between Session 2 and Session 3 was only one day. For that reason, the prototype suffered only a few changes. For this session the rewards were implemented and emphasised the information of lives in the header panel.

It was expected that the participants of this session would be exclusively students from the fourth scholar year. But that was not the case. This time the adaptive interface mode could not be tested as it should because the participants were only allowed to perform one literacy assessment.

5.4.3.1 Detected Anomalies

- ***Level 1*:** the *Level 1* was with no surprise a problem to participants, since it was not possible to improve it. One user took almost 50 seconds to understand what it was requested and about 10 seconds to validate the given answer by the "Check" button. Also, Figure 5.9 shows one of those cases, where even after tapping the answer box and the keyboard appearing, the user tried to tap the underscore character, which hides the missing letter.

Figure 5.9: One user tapping the underscore character of *Level 1*, in Session 3.

- **Reward screen:** when the users levelled up for the first time, the reward screen appeared showing their achievement. To go further the users should tap the "Continue" button. However, a few users tapped first on the middle of the screen to perform that action instead of tapping the button. In Figure 5.10 is it possible to observe one of these users without knowing what to do.

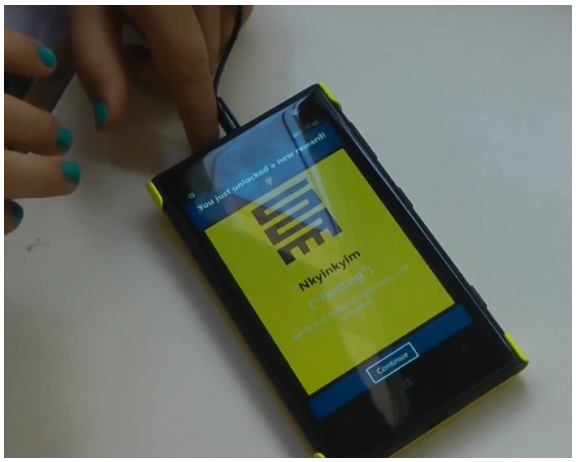


Figure 5.10: One user feeling lost when the reward screen appeared, in Session 3.

5.4.3.2 Additional Observations

Most of the users asked to repeat the experience once the assessment had terminated. Just as it happened in the previous session, it was possible to often listen participants ask the others about their final results. Also, when the evaluation process finished, a group of participants asked to "play" again and see who was the participant with the best classification.

Regarding the open answer questions of the second survey, one user suggested audio feedback when levelling up.

5.4.3.3 Results

The results of the SUS questionnaires for the Session 3 are presented in Table 5.8. Comparatively to the previous sessions, the mean average was the lowest, with a score of 71.79 and classified in the "Ok" zone, near the "Good". It was known that the *Level 1* had already identified usability issues that were not corrected because of the short time interval between the Sessions 2 and 3. Like in the previous sessions, *Level 1* was the level that created more barriers to the users. However, in previous sessions the users easily overcame those problems by the next attempts. In Session 3 that was not the case: almost every users had just one attempt to perform the assessment. That fact can help to explain the decreasing of the SUS score.

Table 5.8: Global SUS score of the Session 3.

Session 3	
User	SUS Global Score
U018	50
U019	82,5
U020	65
U021	52,5
U022	90
U023	95
U024	67,5
Mean	71,79
Standard Deviation	17,78
Minimum	50
Maximum	95

Regarding the second survey, similar results for the questions *Q1* and *Q2* are presented in Figure 5.11. As in the Session 2, aside from a single user (user *U021*), all the users strongly agreed with both questions. Sessions 2 and 3 revealed best results in this questionnaire than Session 1. Maybe the age group of the participants is related, since in the Session 1 the participants were aged between 11 and 12. Nevertheless, the fact is that all the implemented gamification elements only appeared in Session 3 and the results remained satisfactory as well.

In Table 5.9 is shown that although three participants (*ID304*, *ID305*, *ID306*) reached the *Level 4* , none completed the level. In fact, no one managed to reach *Level 5*. Only two users (*ID301* and *ID302*) had equal results in both tests.

Finally, by analysing the JSON files of each assessment,

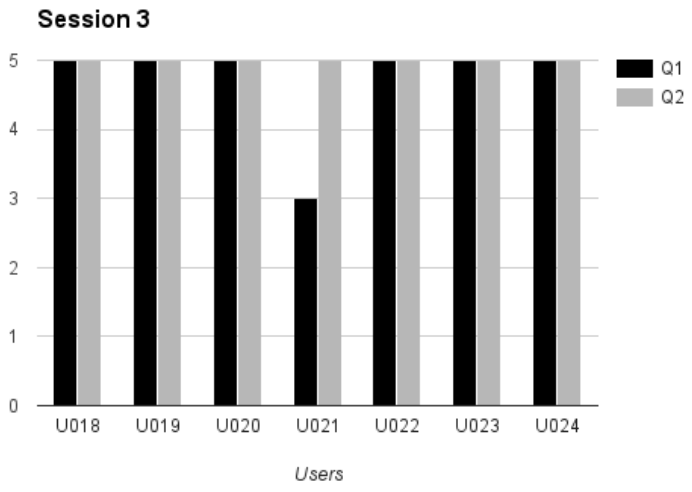


Figure 5.11: Results of the second survey of the Session 3.

Table 5.9: Comparison between the basic literacy test of VPWA and LiteracyTracker, with results of the participants of Session 3.

	SESSION 3				
	Basic Literacy Test			LiteracyTracker	
	Exercise completion				
ID	Words Recognition	Comprehension of sentences	Literacy Profile	Level Reached	Literacy Profile
ID301	yes	no	**can read words**	*Level 3*	**can read words**
ID302	yes	no	**can read words**	*Level 3*	**can read words**
ID303	yes	yes	**can read and understand sentences**	*Level 2*	**can read and write alphabet letters**
ID304	yes	yes	**can read and understand sentences**	*Level 4*	**can read and understand words**
ID305	yes	yes	**can read and understand sentences**	*Level 4*	**can read and understand words**
ID306	yes	yes	**can read and understand sentences**	*Level 4*	**can read and understand words**
ID307	yes	no	**can read words**	*Level 2*	**can read and write alphabet letters**

Tables 5.7 and 5.9 intend to show that by recognizing the sound of some words and pointing them on a piece of paper is not enough to prove and to say that people can actually read words. First because they may simply be guessing, but foremost because it is also important to understand if aside from recognizing the sound of the word they are also capable of associate the word and its meaning, as well as to write what they listen. These components are essential for the everyday life of the target users. Also, besides the fact that some users were categorized as being able to read and understand sentences, on the pen-and-paper based test, LiteracyTracker shows that it might be true, but they still weren't able to correctly write most of the simple words of the *Level 4*.

5.5 Summary

Although it was not possible to evaluate the implemented prototype with the target population, three sessions of tests were conducted in three schools, with 24 participants aged between 7 and 12. After the presentation of the purposes of the evaluation tests, in Section 5.1, details about the sample of participants are given in Section 5.2. The sample is divided into two groups: one composed by 8 mandatory English learners, in the 6th year of school, and another one composed by 16 participants attending the 1st cycle of basic education, learning the English language as their own option.

In Section 5.3, the general procedures of each evaluation session was detailed. Also the surveys used in the evaluation process were explained.

Each evaluation session was described in the Section 5.4. Detected anomalies , observations during the tests and the analysis of the results are also presented in this section.

The gamification of the literacy was well received by the generality of the participants. It keep them engaged with the application and always motivated to improve their results, by starting new assessments. That happened even in a phase that extrinsic rewards were not yet implemented. The adaptive interface showed promising results as well. As the English language was not the first spoken language of the participants, the benefits were not so evident to observe. However, it was observed that mainly for the older participants, with better knowledge of the English language, that the provided explanation of the levels on their mother tongue showed effect.

The analysis of the behaviour of the users, showed also that the command bar of the application was rarely used, since the number of taps in the help button was minimum and only performed by three users.

Also, comparisons between LiteracyTracker and a pen-and-paper based test were made, showing that there a a gap in pen-and-paper based assessments like the one used by the VPWA. They classify users as capable to read even not knowing if they actually understand the meaning of the words. The capability to write simple words is also not taken into account.

Chapter 6

Conclusions

6.1 Final Remarks

An opportunity to overcome a severe problem as illiteracy emerges in the SSA countries, the geographic area with the lowest youth and adult literacy rates in the whole world. On these countries there is not an easy and freely available way to check and assess literacy levels of the populations. The current assessment process often relies on pen-and-paper based tests and is not conducted by specialized people. Thus, since the mobile market in this region is revealing a high growing rate, ICT, namely mobile phones, may take an important role in tracking literacy levels but also in motivating people to learn how to read and write. Additionally, specific behaviours can be identified and analysed, helping to adapt interfaces for each user and mitigating the need for specialized people to conduct the tests.

The literature review showed that it is easy to find numerous studies, tools and technological applications oriented to developing countries that evaluate literacy. However, one major flaw was identified. Most of the applications presented themselves as merely reading tutors, not evaluating the speech or even writing capabilities of the people. Adding to this, those applications do not contemplate any kind of engagement for the user. Gamification and game-based applications are being used in many different areas, namely learning and education, but gamifying literacy is not so usual. From a scientific perspective, this fact increases the need to analyse the effect of these techniques in people's actions in order to persuade them to improve their literacy levels.

In this dissertation a solution was described and a proof of concept was implemented to fulfil those mentioned gaps. LiteracyTracker is the name of the proposed solution, which aims to evaluate the literacy levels of SSA communities by making use of gamification techniques to make assessments enjoyable experiences for the users. The application contains a new literacy test model specially designed after the literature review.

Users face a sequence of levels of exercises that they should complete to succeed and then advance to the next one. Points are earned or lost whether the answer is correct or incorrect. Over time, there are rewards that may be unlocked as the user reaches predetermined goals. Literacy Tracker's interface is adaptive, which means that it changes to the needs of the users. Reducing

visual noise or showing a step by step explanation of the levels goals in the native language of the users, along with audio support, are examples of performed adaptations.

In order to evaluate if gamifying literacy would bring the desired motivation and fun to the participants and to evaluate the feasibility of an adaptive interface, three sessions of tests were conducted with 24 participants. The sample was divided into two groups. The first was composed by 8 mandatory English learners, aged between 11 and 12 and in the 6th year of school. The second group was composed by 16 participants attending the 1st cycle of basic education and aged from 7 to 10, learning the English language as their own option.

The results showed the efficacy of gamifying the literacy in order to provide a captivating experience to the users. It was well received by the generality of the participants. The difficulty level of the exercises used in the evaluation sessions, was a few steps above the knowledge of every participants. Yet, that was not a big barrier to the participant much because of the gamification effects. The presence of the levels and points was decisive to motivate them, keeping them engaged with the application and always motivated to improve their last results by starting new assessments. Also, although no leaderboards were implemented, the score system used in the application has promoted informal and improvised healthy competitions between some of the users, since they revealed ample interest in knowing the classification of each other. That happened even during the first two phases of evaluation, when extrinsic rewards were not yet implemented. That was a good indicator of the feasibility of the designed system. Regarding the adaptive interface, it showed promising results as well. However, as the English language was not the first spoken language of the participants, the benefits were not so evident to observe as they could.

6.2 Future Work

Regarding the prototype system developed, room for improvement has been identified in the four main layers of the system: the literacy assessment model, the gamification of literacy, the HCI guidelines and the adaptive interface mode. Also, as it was expected from a proof of concept, many of the specified features of the solution should be implemented, namely the server side the system.

The designed literacy model revealed potential to be applied and improved. The exercises were validated by an English teacher, however, due to the importance of this subject, it would be useful to ensure that the same model would be validated by more professionals. Apart from that, new kinds of exercises for the existing skills could integrate the designed model, exploring, for example, onset and rime sounds, homophone and homograph words, odd out words, among others.

Not all the specified gamification elements were implemented. Individual and team leaderboards, for communities or countries, would be easy to implement starting from the current state of the application. More than just to addding another gamification element to the list, this feature would bring the social component that is lacking in the actual prototype.

Illiteracy must be opposed within and between communities and its countries. It's a huge problem to solve, essentially in the SSA, and for that reason one of the aims of LiterayTracker is to give its contribution to society. Also, it would open the doors for the *socializers*, the kind of players that are not properly contemplated in this prototype, yet.

New levels can be added with the speech recognition technology. In this prototype only five of the proposed tasks were implemented as levels. Speech recognition would be essential since the application is already exploring the reading, listening and writing skills of the users. However, studies should be made first, to ensure the feasibility and quality of that feature, since if it is incorrectly applied, it could make the users' experience more uncomfortable.

New ways of visual and audio feedback can be explored. For example, along with the visual feedback, different sounds can appear when the user gives correct or incorrect answers. The feedback can be more adjustable to the user in session. By knowing the user's age and gender, for example, different kinds of feedback may be provided.

Regarding the adaptive interface mode, there are many aspects that can still be explored. Machine learning techniques, for example, can be studied to integrate the system, in order to better analyse the behaviours of the users and to give more precise recommendations. Also, it would be useful to test and validate the current implementation of the adaptive mode with users that use English as their first spoken language.

The evaluation results revealed some issues that should be corrected, as well. *Level 1* is clearly the most alarming case. A new solution for its layout should be analysed and then it should be redefined.

Finally, all of this work will only make sense at the moment that LiteracyTracker is tested on the field, as it is expected, by evaluating all the pros and cons of the current prototype and the specified solution, in order to better address the target users needs. New iterations of the evaluation and refinement of design should be made, not only with children but also with adults. It's crucial that these individuals validate many elements of the application, such as icons, colour schemes, navigation structure or used images.

References

[ASH13] Samuel Abramovich, Christian Schunn, and Ross Mitsuo Higashi. Are badges useful in education?: it depends upon the type of badge and expertise of learner. *Educational Technology Research and Development*, 61:217–232, 2013.

[Bad15a] Badgeville.com. Gamification industries & examples, 2015. Available at http://badgeville.com/wiki/Gamification_Industries_Examples.

[Bad15b] Badgeville.com. Gamification psychology, 2015. Available at http://badgeville.com/wiki/psychology.

[BAG14] Kevin Browne, Christopher Anand, and Elizabeth Gosse. Gamification and serious game approaches for adult literacy tablet software. *Entertainment Computing*, 5(3):135–146, August 2014. Available at http://www.sciencedirect.com/science/article/pii/S1875952114000147.

[Bar96] Richard Bartle. Hearts, clubs, diamonds, spades: Players who suit MUDs. *Journal of MUD Research*, 1:19, 1996. Available at http://mud.co.uk/richard/hcds.htm.

[BKM09] Aaron Bangor, Philip Kortum, and James Miller. Determining what individual sus scores mean: Adding an adjective rating scale. *Journal of usability studies*, 4(3):114–123, 2009.

[Blo15] Open Badges Blog. Carla casilli | the myth of the lightweight badge, 2015. Available at http://openbadges.tumblr.com/post/80170437440/carla-casilli-the-myth-of-the-lightweight-badge.

[Bro96] John Brooke. Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189(194):4–7, 1996.

[Chi05] J Chipchase. Understanding non-literacy as a barrier to mobile phone communication. retrieved september 16, 2008, 2005.

[Cod15] Codecademy. Learn to code, 2015. Available at http://www.codecademy.com/.

[Deb11] Educational Technology Debate. Ict and the early grade reading assessment: From testing to teaching, 2011. Available at http://edutechdebate.org/reading-skills-in-primary-schools/ict-and-the-early-grade-reading-assessment-from-testing-to-teaching/

[Des01] Experience Intelligent Design. Cultural Dimensions and Global Web Design : What ? So What ? Now What ? *Design*, page 31, 2001. Available at http://www.amanda.com/resources/hfweb2000/AMA_CultDim.pdf.

[Des14] Service Innovation & Design. Gamification: A quick introduction, 2014. Available at http://sidlaurea.com/2014/09/12/gamification-a-quick-introduction/.

[Det11] Sebastian Deterding. Situated motivational affordances of game elements : A conceptual model. *Chi 2011*, pages 3–6, 2011.

[Duo15] Duolingo.com. Duolingo | learn spanish, french, german, portuguese, italian and english for free, 2015. Available at https://www.duolingo.com/.

[FBT09] Leah Findlater, Ravin Balakrishnan, and Kentaro Toyama. Comparing semiliterate and illiterate users' ability to transition from audio+ text to text-only interaction. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, page 1751, 2009. Available at http://dl.acm.org/citation.cfm?doid=1518701.1518971$\delimiter"026E30F$nhttp://dl.acm.org/citation.cfm?id=1518971.

[Fog09] Bj Fogg. A behavior model for persuasive design. *Proceedings of the 4th International Conference on Persuasive Technology - Persuasive '09*, page 1, 2009. Available at http://portal.acm.org/citation.cfm?doid=1541948.1541999.

[fS] UNESCO Institute for Statistics. *The next generation of literacy statistics*. Number 1.

[Gal11] Scott Gallacher. Jon Radoff – Game On: Energize Your Business with Social Media Games, 2011.

[Gav12] Takayedzwa Gavaza. *Culturally-relevant Augmented User Interfaces for Illiterate and Semi-literate Users*. PhD thesis, Rhodes University, 2012.

[Gha15] Ghanaembassy.nl. What is the official language of ghana?, 2015. Available at http://www.ghanaembassy.nl/index.php/faqs-mainmenu-25/121-what-is-the-official-language-of-ghana.html.

[Gon13] Goncalo Goncalves. *Promover o Turismo em Portugal com recurso a Gamification*. PhD thesis, Faculdade de Engenharia da Universidade do Porto, 2013.

[Gov08] A Gove. Early grade reading assessment toolkit. *World Bank and RTI International, Washington, DC*, 2008.

[GSM14] GSMA. Latest report from gsma, 2014. Available at http://www.gsmamobileeconomyafrica.com/.

[Gsm15a] Gsmarena.com. Nokia lumia 1020 - full phone specifications, 2015. Available at http://www.gsmarena.com/nokia_lumia_1020-5506.php.

[Gsm15b] Gsmarena.com. Nokia lumia 630 - full phone specifications, 2015. Available at http://www.gsmarena.com/nokia_lumia_630-6232.php.

[Gsm15c] Gsmarena.com. Nokia lumia 830 - full phone specifications, 2015. Available at http://www.gsmarena.com/nokia_lumia_830-6638.php.

[Gsm15d] Gsmarena.com. Nokia lumia 920 - full phone specifications, 2015. Available at http://www.gsmarena.com/nokia_lumia_920-4967.php.

[Hag12] Per Hagglund. *Taking gamification to the next level A detailed overview of the past, the present and a possible future of gamification*. PhD thesis, Umeå University, 2012.

[Hue08] Matthew P Huenerfauth. Developing design recommendations for computer interfaces accessible to illiterate users. 2008.

[Inf] Info.graphogame.com. Graphogame | evidence-based learning game helps children to learn to read. Available at http://info.graphogame.com/.

[JWKS08] Anirudha Joshi, Nikhil Welankar, Kirti Kanitkar, and Riyaj Sheikh. Rangoli: a visual phonebook for low-literate users. In *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*, pages 217–223. ACM, 2008.

[KAK+08] Matthew Kam, Aishvarya Agarwal, Anuj Kumar, Siddhartha Lal, Akhil Mathur, and John Canny. Designing E-Learning Games for Rural Children in India : A Format for Balancing Learning with Fun. *Electrical Engineering*, pages 58–67, 2008. Available at http://portal.acm.org/citation.cfm?doid=1394445.1394452.

[Kap15] Karl Kapp. Measurement vs. completion achievements in gamification « karl kapp, 2015. Available at http://karlkapp.com/measurement-vs-completion-achievements-in-gamification/.

[Kou15] Mawuna Koutonin. 63 african symbols for creative design, 2015. Available at http://www.siliconafrica.com/african-symbols-for-creative-design/.

[KRD+07] Matthew Kam, Divya Ramachandran, Varun Devanathan, Anuj Tewari, and John Canny. Localized iterative design for language learning in underdeveloped regions: the PACE framework. In *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '07*, page 1097, 2007. Available at http://portal.acm.org/citation.cfm?id=1240791$\delimiter"026E30F$nhttp://portal.acm.org/citation.cfm?doid=1240624.1240791.

[KRT+12] Anuj Kumar, Pooja Reddy, Anuj Tewari, Rajat Agrawal, and Matthew Kam. Improving literacy in developing countries using speech recognition-supported games on mobile devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1149–1158. ACM, 2012.

[Kuu13] Julius Kuutti. *DESIGNING GAMIFICATION*. PhD thesis, OULU BUSINESS SCHOOL, 2013.

[Kwi15] Kwintessential.co.uk. Ghana - language, culture, customs and etiquette | global-etiquette | resources, 2015. Available at http://www.kwintessential.co.uk/resources/global-etiquette/ghana.html.

REFERENCES

[Laz04] N. Lazzaro. Why We Play Games: Four Keys to More Emotion Without Story. Technical report, 2004. Available at http://www.xeodesign.com/xeodesign_whyweplaygames.pdf.

[Lel14] Clemence Lelong. Gamification design framework (from professor werbach), 2014. Available at https://clemencelelong.wordpress.com/2014/01/29/gamification-design-framework-from-professor-werbach/.

[LG08] Zereh Lalji and Judith Good. Designing new technologies for illiterate populations: A study in mobile phone interface design. *Interacting with computers*, 20(6):574–586, 2008.

[Mac15] Jean MacDonald. Adinkra symbols of west africa, 2015. Available at http://www.adinkra.org/.

[MMCT10] Indrani Medhi, S Raghu Menon, Edward Cutrell, and Kentaro Toyama. Beyond strict illiteracy: abstracted learning among low-literate users. In *Proceedings of the 4th ACM/IEEE International Conference on Information and Communication Technologies and Development*, page 23. ACM, 2010.

[MPT05] Indrani Medhi, Bharathi Pitti, and Kentaro Toyama. Text-free user interfaces for employment search. In *Asian Applied Computing Conference, Nepal*. World Scientific, 2005.

[MPT07] Indrani Medhi, Archana Prasad, and Kentaro Toyama. Optimal audio-visual representations for illiterate users of computers. In *Proceedings of the 16th international conference on World Wide Web*, pages 873–882. ACM, 2007.

[Msd15] Msdn.microsoft.com. Onscreen keyboard input scope index for windows phone 8, 2015. Available at https://msdn.microsoft.com/en-us/library/windows/apps/hh393998(v=vs.105).aspx.

[MST06] Indrani Medhi, Aman Sagar, and Kentaro Toyama. Text-free user interfaces for illiterate and semi-literate users. In *Information and Communication Technologies and Development, 2006. ICTD'06. International Conference on*, pages 72–82. IEEE, 2006.

[Nor13] Donald A Norman. *The design of everyday things: Revised and expanded edition*. Basic books, 2013.

[Ogd44] Charles Kay Ogden. *Basic English: A general introduction with rules and grammar*. Number 29. K. Paul, Trench, Trubner, 1944.

[ohtkotYcla15] You only have to know one thing: You can learn anything. Khan academy, 2015. Available at https://www.khanacademy.org/.

[PGC03] Tapan Parikh, Kaushik Ghosh, and Apala Chavan. Design studies for a financial management system for micro-credit groups in rural india. In *ACM SIGCAPH Computers and the Physically Handicapped*, number 73-74, pages 15–22. ACM, 2003.

REFERENCES

[PMTB08] Archana Prasad, Indrani Medhi, Kentaro Toyama, and Ravin Balakrishnan. Exploring the feasibility of video mail for illiterate users. In *Proceedings of the working conference on Advanced visual interfaces*, pages 103–110. ACM, 2008.

[Por] Fraunhofer Portugal. Fraunhofer portugal - about us. Available at http://www.fraunhofer.pt/en/fraunhofer_portugal/about_us.html.

[Por15] Fraunhofer Portugal. Acp street libraries, 2015. Available at http://www.fraunhofer.pt/en/fraunhofer_portugal/news/news_archive/acp-street-libraries.html.

[Qui15] Leonard Ellom Quist. The adinkra symbols and icons project, 2015. Available at http://adinkraproject.com/.

[RKTL14] Miia Ronimus, Janne Kujala, Asko Tolvanen, and Heikki Lyytinen. Children's engagement during digital game-based learning of reading: The effects of time, rewards, and challenge. *Computers and Education*, 71:237–246, 2014. Available at http://dx.doi.org/10.1016/j.compedu.2013.10.008.

[Rug13] Razvan Rughinis. Gamification for Productive Interaction Reading and Working with the Gamification Debate in Education. In *The 8th Iberian Conference on Information Systems and Technologies CISTI 2013*, pages 1–5, 2013.

[Rya09] Richard M. Ryan. Self - determination Theory and Wellbeing. *Wellbeing in Developing Countries*, pages 1–2, 2009.

[San15] Joana Santos. Sistemas de ensino na união europeia, 2015. Available at http://www.educare.pt/noticias/noticia/ver/?id=13150.

[Sau15] Jeff Sauro. Measuring usability with the system usability scale (sus): Measuringu, 2015. Available at http://www.measuringu.com/sus.php.

[Sch08] Jesse Schell. *The Art of Game Design: A Book of Lenses*, volume 54. 2008. Available at http://books.google.com/books?hl=en&lr=&id=LP5xOYMjQKQC&oi=fnd&pg=PP1&dq=The+Art+of+Game+Design:+A+Book+of+Lenses&ots=Y96OsGL9yI&sig=Atib_ofH5sSvrlUg4xVayecQ_vc.

[SPM+09] Jahanzeb Sherwani, Sooraj Palijo, Sarwat Mirza, Tanveer Ahmed, Nosheen Ali, and Roni Rosenfeld. Speech vs. touch-tone: Telephony interfaces for information access by low literate users. In *Information and Communication Technologies and Development (ICTD), 2009 International Conference on*, pages 447–457. IEEE, 2009.

[SU10] United Nations Educational Scientific and Cultural Organization (UNESCO). *Education for All Global Monitoring Report 2010: Reaching the marginalized (EFA Global Monitoring Report)*. Oxford University Press, 2010. Available at http://unesdoc.unesco.org/images/0018/001866/186606E.pdf.

[Tan] Tangerinecentral.org. About tangerine | tangerine central. Available at http://www.tangerinecentral.org/aboutsite.

[Ter] Bénédicte Terryn. Measuring Literacy in Developing Countries from an International Perspective. *Statistics*.

[Tet06] AV Tetteh. Adinkra-cultural symbols of the asante people. Technical report, Retrieved 07/09/12, from, 2006.

[Tew15] Anuj Tewari. Language-learning games - millee, 2015. Available at http://www.cs.berkeley.edu/~anuj/language_learning_games/MILLEE.html.

[tL15] Project LISTEN A Reading Tutor that Listens. Summary, 2015. Available at http://www.cs.cmu.edu/~./listen/research.html.

[TLW+10] Feng Tian, Fei Lv, Jingtao Wang, Hongan Wang, Wencan Luo, Matthew Kam, Vidya Setlur, Guozhong Dai, and John Canny. Let's Play Chinese Characters: Mobile Learning Approaches via Culturally Inspired Group Games. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 1603–1612, New York, NY, USA, 2010. ACM. Available at http://doi.acm.org/10.1145/1753326.1753565.

[TS] Jamie Thomas and Catherine Stringfellow. Using HCI Guidelines to Foster Technological Advancement in Sub-Saharan Africa. 2.

[UNE12] University of La Laguna UNESCO. *Expert Meeting on Early Grades Writing Assessment (EGWA)*. Expert Meeting on Formative Assessment of Writing in Early Grades in Primary Education. 2012. Available at http://unesdoc.unesco.org/images/0022/002209/220903E.pdf.

[Une15] Unesco.org. Writing assessment in early grades of primary school (egwa) | education | united nations educational, scientific and cultural organization, 2015. Available at http://www.unesco.org/new/en/education/themes/education-building-blocks/literacy/writing-assessment-in-early-grades-of-primary-school-egwa/.

[Usa15a] Usabilityfirst.com. Usability first - introduction to user-centered design | usability first, 2015. Available at http://www.usabilityfirst.com/about-usability/introduction-to-user-centered-design.

[Usa15b] Usability.gov. System usability scale (sus), 2015. Available at http://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html.

[vdB14] Chris van den Berg. *Gamify a Contact Center*. PhD thesis, University of Amsterdam, 2014.

[VG12] R Vesselinov and J Grego. Duolingo Effectiveness Study. *City University of New York, USA*, 2012. Available at http://static.duolingo.com/s3/DuolingoReport_Final.pdf.

[Wag03] Daniel A. Wagner. Smaller, quicker, cheaper: Alternative strategies for literacy assessment in the UN Literacy Decade. *International Journal of Educational Research*, 39:293–309, 2003.

[WH12] Kevin Werbach and Dan Hunter. *For the win*. Wharton Digital Press, 2012.

REFERENCES

[Wu11] Michael Wu. Gamification from a company of pro gamers, 2011. Available at http://lithosphere.lithium.com/t5/Building-Community-thePlatform/Gamification-from-a-Company-of-Pro-Gamers/ba-p/19258.

[ZC08] G. Zichermann and C. Cunningham. *Gamification By Design*. 2008. Available at http://medcontent.metapress.com/index/A65RM03P4874243N.pdf.

REFERENCES

Appendix A

Configuration Files

A.1 XML Levels Database Example

In this example, it is presented the XML structure of the *Level 4*. Each database file starts equally with the definition of the configuration elements, such as the level id, level title or the textual and audio help that will support the level, for example. Next, the second part of the file defines the content of the exercises. In this case, the exercise requires a list of words, so each one of them is inserted into the file following the seen in Figure A.1. In this example, only a few words are defined, however, if that was the original file, easily the list of words could be increased. All the attributes of the "*LevelConfiguration*" tag can also be changed. For example, to update de number of rounds of the level, it only needed to change the attribute "rounds" for another number. The application adapts to the changes of these types of file.

```
<?xml version="1.0" encoding="utf-8" ?>
<Level4>
  <Configs>
    <!--native_lang is the native language of the user. It will be important, to let output tips in that same language-->
    <LevelConfiguration id="4" title="Level 4" rounds="6" lives_left="3" correct_points="180" incorrect_points="75" lang="en-GB" native_lang="none"/>
    <Description text=""/>
    <TextHelp text="Write the word that you hear."/>
    <AudioHelp text="Write the word that you hear."/>
  </Configs>
  <ListofWords>
    <word>account</word>
    <word>act</word>
    <word>addition</word>
    <word>adjustment</word>
    <word>advertisement</word>
    <word>agreement</word>
    <word>air</word>
    <word>amount</word>
    <word>amusement</word>
    <word>animal</word>
    <word>answer</word>
    <word>apparatus</word>
    <word>approval</word>
    <word>argument</word>
    <word>art</word>
    <word>attack</word>
    <word>attempt</word>
    <word>attention</word>
  </ListofWords>
</Level4>
```

Figure A.1: Example of the structure of *Level 4* database.

A.2 JSON Files Examples

Figure A.2 shows an extract of a JSON file generated by LiteracyTracker in the end of each assessment. In this example, is possible to see that in the *Level 3* the user incorrectly matched the word "pencil" with the image "cow" and the word "heart" with the image "hat". Moreover, it is possible to see the earned points, each level duration or each level staring date, for instance.


```
{
  "WIN_LEVEL3_REWARD": 3,
  "LevelTitle": "Level 3",
  "GeneratedSolutions": [
    "cup",
    "sheep",
    "needle",
    "pencil",
    "horse",
    "heart"
  ],
  "TappedImages": [
    "cup",
    "sheep",
    "needle",
    "cow",
    "horse",
    "hat"
  ],
  "LevelSolution": "heart",
  "IdLevel": 3,
  "LevelDifficulty": 1,
  "RoundsLeft": 0,
  "LevelStartDate": "11/06/2015 18:14:35",
  "LevelCompleted": true,
  "LevelDurationTime": "32.3825125",
  "LevelPointsEarned": 915,
  "LevelPointsRightAnswer": 110,
  "LevelPointsWrongAnswer": 75,
  "RightAnswersCount": 4,
  "HasRewardElement": false
},
```

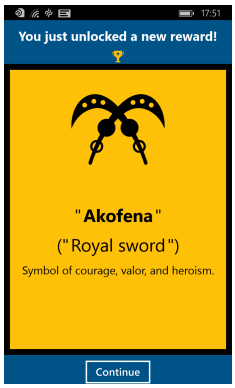

Figure A.2: Example of a JSON file generated in the end of each assessment test.

Appendix B

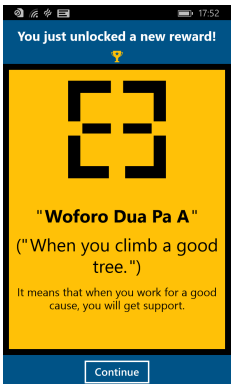
Additional Screens

B.1 Unlocked Rewards Examples

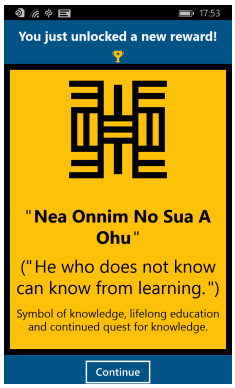
(a) Example of unlocked reward(i).



(b) Example of unlocked reward(ii).



(c) Example of unlocked reward(iii)

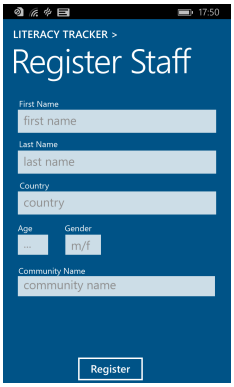


(d) Example of unlocked reward(iv)

Figure B.1: More examples of unlocked rewards of the application.

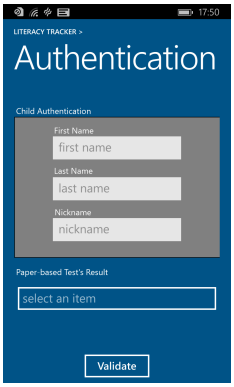
B.2 Additional Screens Examples

The following are screens of the registration and authentication process of the application.

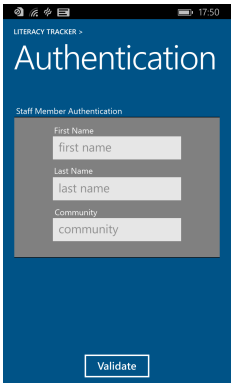


(a) Registration of Staff Members.

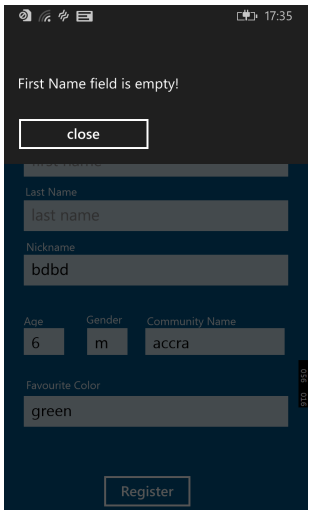
(b) Registration of Regular Users.



(c) Regular User authentication.



(d) Staff authentication.



(e) Example validation message, when a field is empty.

Figure B.2: Additional screen about the registration and the authentication process.

Appendix C

Questionnaires

C.1 SUS Questionnaire

1. Questionário SUS (System Usability Scale Questionnaire)

Usando a escala abaixo, por favor coloque um círculo no número mais próximo da palavra que mais se aproxima aos seus sentimentos acerca da aplicação.

1. **Penso que gostaria de usar este sistema frequentemente.**

Discordo fortemente	1	2	3	4	5	Concordo fortemente

2. **Achei o sistema desnecessariamente complexo.**

Discordo fortemente	1	2	3	4	5	Concordo fortemente

3. **Achei o sistema fácil de usar.**

Discordo fortemente	1	2	3	4	5	Concordo fortemente

4. **Penso que precisaria do apoio técnico para conseguir usar o sistema.**

Discordo fortemente	1	2	3	4	5	Concordo fortemente

5. **Achei que as várias funções do sistema estavam bem integradas.**

Discordo fortemente	1	2	3	4	5	Concordo fortemente

6. **Achei que havia demasiadas inconsistências neste sistema.**

Discordo fortemente	1	2	3	4	5	Concordo fortemente

7. **Imagino que a maioria das pessoas consegue aprender a usar este sistema muito rapidamente.**

Discordo fortemente	1	2	3	4	5	Concordo fortemente

8. **Achei o sistema muito incómodo de usar.**

Discordo fortemente	1	2	3	4	5	Concordo fortemente

9. **Senti-me muito confiante ao usar o sistema.**

Discordo fortemente	1	2	3	4	5	Concordo fortemente

10. **Precisei de aprender muitas coisas antes de conseguir começar a usar o sistema.**

Discordo fortemente	1	2	3	4	5	Concordo fortemente

Figure C.1: SUS questionnaire used in the evaluation sessions(in Portuguese).

C.2 Second Survey

2. Questionário Suplementar

1. Achei interessante a diferente gama de exercícios apresentados.

Discordo fortemente	1	2	3	4	5	Concordo fortemente

2. Achei interessante o sistema de pontos e vidas por níveis.

Discordo fortemente	1	2	3	4	5	Concordo fortemente

3. Críticas, dificuldades, sugestões ou outros comentários adicionais (opcional) :

Figure C.2: Second questionnaire used in the evaluation sessions(in Portuguese).

C.3 SUS Score of Sessions

C.3.1 Session 1

Table C.1: SUS score calculation of the Session 1.

		U001	U002	U003	U004	U005	U006	U007	U008
Raw Scores	Q1	4	4	4	4	5	5	5	4
	Q2	1	3	1	1	1	1	1	1
	Q3	5	5	5	5	5	5	5	5
	Q4	1	1	1	1	1	1	1	1
	Q5	5	4	4	5	1	5	5	5
	Q6	1	3	1	1	1	1	1	2
	Q7	5	5	4	4	5	4	5	5
	Q8	1	1	4	1	1	1	1	1
	Q9	5	4	4	5	5	5	5	5
	Q10	1	1	1	1	1	3	2	1
Calculated Scores	Q1	3	3	3	3	4	4	4	3
	Q2	4	2	4	4	4	4	4	4
	Q3	4	4	4	4	4	4	4	4
	Q4	4	4	4	4	4	4	4	4
	Q5	4	3	3	4	0	4	4	4
	Q6	4	2	4	4	4	4	4	3
	Q7	4	4	3	3	4	3	4	4
	Q8	4	4	1	4	4	4	4	4
	Q9	4	3	3	4	4	4	4	4
	Q10	4	4	4	4	4	2	3	4
Total Score		39	33	33	38	36	37	39	38
SUS Global Score		97,5	82,5	82,5	95	90	92,5	97,5	95
Average SUS Global Score		91,56							

C.3.2 Session 2

Table C.2: SUS score calculation of the Session 2.

		U009	U010	U011	U012	U013	U014	U015	U016	U017
Raw Scores	Q1	5	5	5	5	5	5	5	5	4
	Q2	1	1	1	1	3	1	1	1	3
	Q3	3	5	5	3	3	5	5	5	4
	Q4	3	1	4	3	1	1	4	1	4
	Q5	3	5	5	5	5	5	5	3	2
	Q6	1	1	1	1	1	1	1	5	5
	Q7	5	1	4	3	5	4	1	5	3
	Q8	4	1	1	1	1	1	1	5	1
	Q9	5	5	5	5	5	5	5	5	4
	Q10	5	5	5	3	5	1	3	5	5
Calculated Scores	Q1	4	4	4	4	4	4	4	4	3
	Q2	4	4	4	4	2	4	4	4	2
	Q3	2	4	4	2	2	4	4	4	3
	Q4	2	4	1	2	4	4	1	4	1
	Q5	2	4	4	4	4	4	4	2	1
	Q6	4	4	4	4	4	4	4	0	0
	Q7	4	0	3	2	4	3	0	4	2
	Q8	1	4	4	4	4	4	4	0	4
	Q9	4	4	4	4	4	4	4	4	3
	Q10	0	0	0	2	0	4	2	0	0
Total Score		27	32	32	32	32	39	31	26	19
SUS Global Score		67,5	80	80	80	80	97,5	77,5	65	47,5
Average SUS Global Score		75,00								

C.3.3 Session 3

Table C.3: SUS score calculation of the Session 3.

		U018	U019	U020	U021	U022	U023	U024
Raw Scores	Q1	5	4	5	5	5	5	5
	Q2	1	1	3	3	1	1	3
	Q3	1	3	5	4	5	5	5
	Q4	5	1	4	5	1	1	4
	Q5	5	3	3	4	3	4	3
	Q6	5	1	4	4	3	1	4
	Q7	1	5	3	5	5	4	5
	Q8	1	1	1	4	1	1	1
	Q9	5	5	5	2	5	5	5
	Q10	5	3	3	3	1	1	4
Calculated Scores	Q1	4	3	4	4	4	4	4
	Q2	4	4	2	2	4	4	2
	Q3	0	2	4	3	4	4	4
	Q4	0	4	1	0	4	4	1
	Q5	4	2	2	3	2	3	2
	Q6	0	4	1	1	2	4	1
	Q7	0	4	2	4	4	3	4
	Q8	4	4	4	1	4	4	4
	Q9	4	4	4	1	4	4	4
	Q10	0	2	2	2	4	4	1
Total Score		20	33	26	21	36	38	27
SUS Global Score		50	82,5	65	52,5	90	95	67,5
Average SUS Global Score		71,79						