FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Smartphone serious games for senses evaluation

Pedro Dias

MESTRADO INTEGRADO EM ENGENHARIA INFORMÁTICA E COMPUTAÇÃO

Supervision: A. Augusto de Sousa

Proposal: Bruno Aguiar, Fraunhofer Portugal

July 28th, 2014

Smartphone serious games for senses evaluation

Pedro Dias

Mestrado Integrado em Engenharia Informática e Computação

Approved in oral examination by the committee:

Chair: Alexandre Miguel Barbosa Valle de Carvalho, PhD
External Examiner: Hugo Alexandre Paredes Guedes da Silva, PhD
Supervisor: António Augusto de Sousa, PhD

July 28th, 2014

Abstract

Falls are the second leading cause of accidental injury deaths worldwide. There is a Project at Fraunhofer which attempts to determine one's fall risk by analyzing a number of metrics obtained through the use of common devices, such as smartphones.

In this dissertation, it is intended to define methodologies that permit the evaluation of two potential factors which might have an impact on fall risk [1]–[3], these are: visual and hearing loss. The aim of the work developed is not to replace clinic visits, but to offer the user the means to continue the tracking of his vision and audition at home, during the long time intervals between clinical tests. If these evaluations are successful, alongside other metrics not covered in this dissertation, it might allow for an assessment of fall risk, so that if dangerous readings are detected the user can be immediately directed to the clinic, thus reducing the risk of serious injury.

To do this, methodologies were proposed that, through the use of an Android application, are capable of evaluating the visual and hearing capabilities of the user, without the need of another person's help. Furthermore, a critical factor in this evaluation is the high frequency of the tests, which increases the data gathering rate and likelihood of problem detection. With the purpose of encouraging their use, the tests were developed as serious games in order to increase the enjoyment of senior users. These games are designed to be as simple, pleasant and intuitive as possible, so that the user will have additional motivation to carry on the tests, allowing the data to be easily and frequently gathered.

Testing conducted in a sample of our target users seemed to indicate a good ability to measure vision and audition using the methodologies developed. While some tests require further validation, promising results were achieved in the most used tests for vision and audition, presenting a good correlation between the system's results when compared to the traditional test (for distance visual acuity) and the data gathered from the users (for audition tests).

The goals for this dissertation are considered fulfilled. The methodologies developed allow the evaluation of vision and audition, without the need of a complex user experience, external help or hard-to-obtain additional hardware.

Acknowledgements

I want to thank both my supervisors, António Augusto de Sousa and Bruno Aguiar, for the invaluable time and advice given throughout all this process. Their constant willingness to help enrich this dissertation, paved the way for a great experience.

I would also like to thank Fraunhofer Portugal for their permanent support and all the resources they made available to me, so that I could make the most of my abilities. All the assistance in developing and testing the work developed here was a key factor in its success. Likewise, I would like to thank the volunteers who were willing to test the prototype system and, therefore, contributed to improvements on the work developed.

Finally, I would like to thank my family and friends for all the support given, not only during this dissertation's development process, but throughout all my college years.

Pedro Dias

Index

List of Figures

List of Tables

Chapter 1

Introduction

Falls are a painful reality for a large percentage of the population, especially senior citizens. It is a common problem that comes with old age and that presents many problems to those afflicted by it, including: injuries, fear, loss of independence and, therefore, a decrease in quality of life. Thus, being able to evaluate one's fall risk can help to prevent such problems and speed up the prevention process [4]. Fraunhofer Portugal has a running project that attempts to make this evaluation using many different metrics acquired via smartphone. Two of these metrics are vision and audition, which are the target senses of this dissertation.

This dissertation focus on developing and testing methodologies that allow the evaluation of the user's vision and audition, through the use of a smartphone based system. In order to be able to target the most affected group (senior citizens) this system must be simple, entertaining and inexpensive. To achieve this, the system focuses on the execution of game-based tests and, at all times, aims for simple interfaces and interactions that are easily operated, even by inexperienced users. For clarity purposes, the game-based tests implemented will be referred to as "exams".

It is hoped that, with such a system in place, valuable audition and vision metrics can be obtained. These might be later used in fall risk evaluation systems, but that is not the goal of this dissertation.

1.1 Motivation and Objectives

As mentioned before, the two senses intended to be evaluated in this dissertation are the users' audition and vision. These two factors are believed to be able to help in the detection of fall risk. Being able to evaluate both senses and detect flaws in one or both of them, alongside others metrics not analyzed in this dissertation, would go a long way towards the final objective of Fraunhofer Portugal's project of estimating fall risk assessment. The exams should be self-administered, without the need of any external help, through serious games developed for Android devices.

The research conducted showed some studies and tools that target similar goals, but are ultimately different and don't meet the requirements of this dissertation. Most of them are not for self-use, instead targeting the professional who will perform the tests using a smartphone (or tablet). Others are very limited in the sense characteristic they evaluate.

Therefore, the research that was made aims to improve this situation, by analyzing and developing a methodology that would provide the user with a way to test two of the most important information channels that people use permanently, and does so without being excessively complex or tedious. Additionally, by giving the user game based tests that are intuitive and easy to use, the likelihood of tainting the results obtained is decreased; and by providing an easy way for the user to conduct them on themselves, without the need of other person's help (ex: Nurse or Care giver), exam frequency is potentially increased.

In order to complete this dissertation, several steps were taken:

1. State of the Art and Specialized consultancy
2. Development of Vision evaluation exams and related games
3. Development of Audition evaluation exams and related games
4. Final User Tests
5. Result Evaluation

Step 1 provided valuable information on which tests are conducted in the current clinical setting, as well as similar works and their advantages and limitations. It provided several considerations to be taken into account when selecting how and which tests should be implemented.

Steps 2 and 3 focused on the development process of the Vision and Audition Evaluation exams selected, respectively. After the system was ready to perform the tasks required, games would be developed for those features.

In Step 4, data was gathered through the use of the system by senior citizens. This is a sample of the target audience and provided valuable considerations on the work developed.

In Step 5, the data gathered in the previous step was analyzed in order to make an evaluation of the test results and to identify areas that could possibly be improved.

1.2 Report's Structure

Besides this introductory chapter, this report also comprises four additional chapters:

- Chapter 2 focuses on the state of the art, presenting the research of papers, tools and clinical tests that were found and deemed relevant to be featured in this report.

- Chapter 3 focuses on the work developed. It includes considerations about designing and developing the methodologies used, explains which exams were implemented and why, and gives a more detailed look at the decisions made during the exams development and their progress.
- Chapter 4 focuses on documenting the tests performed on volunteers, in an attempt to validate the work developed. Details on how the tests were conducted are given; the results obtained are presented and conclusions are made.
- Chapter 5 is the conclusion of this report and summarizes the work developed throughout this dissertation. In this chapter, future work is also mentioned, including possible improvements than can be made and a possible future for the research developed.

Afterwards, a list of references used in this document and Appendices related to the work developed are also listed.

Chapter 2

Related Work

In this chapter, documentation that was found and deemed relevant is analyzed so that a better understanding of what already exists in this area and what can be used as support can be achieved.

A thorough investigation of which tests are used for audition and vision assessment was conducted. Specialists in both areas were contacted and pointed out the tests that are more frequent in their respective fields. Then, both technological and non-technological related work was researched, finding hits on both terms. Similar works are shown and pertinent information is drawn from them.

Some of the information that is considered required for the development of this research is also presented, as well as an explanation of how some of the relevant tests work.

2.1 Vision and Audition Evaluation

This section will focus on traditional tests, conducted in the vision and audition assessment field, that were relevant for the system design phase, with a brief explanation of each one. This will focus on identifying each test's goal, procedures and important considerations.

2.1.1 Vision evaluation

There are a number of tests that can be used to measure vision quality, in terms of visual acuity, visual field, shape discrimination, depth perception or color discrimination. Based on the research conducted, the ones which are used to a higher degree in medical settings and the ones which would have the most benefit from a conversion to a smartphone setting are presented here.

Some of the usual vision related tests, that are believed to be helpful in the evaluation of vision quality, will now be covered.

2.1.1.1 Snellen Chart

This is one of the most common tool which can be found for vision measurement. [5], [6] It received its name from Hermann Snellen who developed it 1862. As shown in Figure 1, this test presents the patient with a series of lines with random letters of decreasing sizes and its purpose is to measure visual acuity, which is seen by many as the most important indicator of vision quality. [7] By finding the smallest line which the patient can read, factoring in the distance between him and the chart, one can assess the visual acuity of the patient. The recommended distance for this test varies between charts, but is normally set between 3 and 6 meters [7]–[9]. One can perform this test at smaller distances by factoring this in the result calculation, or through the use of mirrors to double the perceived distance between user and chart. [10]–[14]

Figure 1. Snellen Chart

Snellen’s work has since been used as the basis for many alternatives which aim to improve on it in different ways, such as: ETDRS Eye Charts [9], [11], Bailey-Lovie Chart [7], [15], Pelli-Robson Contrast Sensitivity Chart [16], [17] and Freiburg Visual Acuity and Contrast Sensitivity Test [18], [19].

Baring certain factors (for example: inability to read the Latin alphabet), different optotypes (i.e. symbols used in vision tests) might be used. Examples of these are: Landolt C’s [18], [20] (Figure 2a), which consists of a ring with a gap at different positions, normally top, bottom, left and right, but that may also be positioned diagonally; Tumbling E’s [21] (Figure 2b), which is an optotype similar to the letter “E”, facing various directions (top, bottom left and right); and Lea symbols [22]–[24], which are a group of symbols, normally used for testing on children.

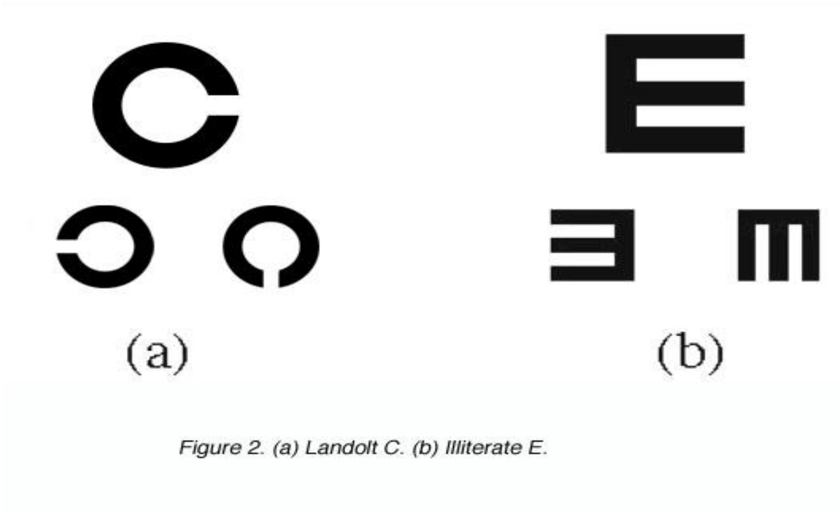

Figure 2. Two different types of optotypes: (a) Landolt's C - (b) Tumbling E

Yet another set of optotypes are the Sloan letters, a combination of the letters C, D, H, K, N, O, R, S, V and Z, designed by Louise Sloan in 1959 as an alternative to Snellen letters. [7], [25] They are now used in Snellen, LogMAR and Pelli-Robson Charts (see following sections). [26]

2.1.1.2 Freiburg Vision Test

The Freiburg Acuity and Contrast Test (FrACT) was developed by Professor Michael Bach and is an automated procedure for self-administered measurement of visual acuity and contrast sensitivity. [12], [18], [19] The test has gone through different iterations and an online example can be found. [27] It offers two measurements, both Visual Acuity and Contrast Sensitivity. Visual Acuity (VA) is measure by presenting either a Landolt C or Tumbling E and asking the user to identify the orientation of the optotype. As the user correctly answers the trials, he will be given increasingly smaller optotypes until his VA is found. The Contrast Sensitivity test works in a similar way, but instead of changing the optotype size, its contrast with the background is altered, thus attempting to find the threshold at which the user can no longer identify the correct orientation. Both of these tests use a best PEST procedure [28], [29], in which the systems alters the step at which the challenge difficulty is increased or decreased, attempting to get closer and closer to the user's true result. This tool also features tests for Vernier Acuity and Contrast Grating, although the documentation is less comprehensive on these tests.

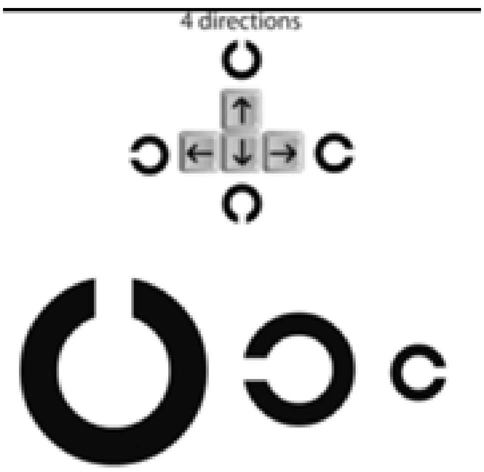

Figure 3. Freiburg Acuity and Contrast Test

2.1.1.3 LogMAR Chart

LogMAR stands for "Logarithm of the Minimum Angle of Resolution". It is a chart similar to the Snellen Chart, which focuses on visual acuity measurement, but it features a constant number of letters per line and uniform progression of letter size and spacing between letters.

Figure 4. LogMAR Chart

It was designed to be used at a 4 meter distance (against Snellen's 6 meters) [30]–[32], thus allowing the test to be taken more easily. Each letter has an equal score value of 0.02 log units, and the value of a line is equal to the number of letters read in that line, since all lines have five letters the maximum and minimum score for a line is 0.1 and 0 respectively. The formula to calculate one's LogMAR Visual Acuity (LogMAR VA) score is [32], [33]:

$$LogMAR\ VA = LogMARb - [0.02 * letterN] \qquad (1)$$

Where $LogMARb$ represents the LogMAR value of the best line completely read and $letterN$ represents the number of letters read in the next line.

A LogMAR score of 0.0 equals 6/6 (1.0 in decimal format) by Snellen standards (i.e. normal vision) and 1.0 equals 6/60 (0.1 in decimal format).

The results of a LogMAR test can then be converted into Snellen acuity. [34]

2.1.1.4 Pelli-Robson Contrast Sensitivity Chart

Unlike the previous charts which focus on visual acuity, this one's purpose is to evaluate one's contrast sensitivity. [20], [35] There are different lines of six letters each, organized in two groups of three (triplets) per line, being the contrast decreased from triplet to triplet. The chart is normally placed one meter away from the patient, who identifies the least contrasted triplet which he can read. [16] One study concluded that the loss of contrast sensitivity can lead to balance and mobility problems, [36] therefore this test could prove itself valuable for fall risk calculation.

Figure 5. Pelli-Robson Contrast Sensitivity Chart

2.1.1.5 Jaeger Eye Chart

This test is intended for near vision measurement, normally to evaluate reading capabilities. The chart consists of blocks of text, with letter size variating between blocks, each one with an associated Jaeger scores (J1, J2, etc.) which can be used to make an estimative of the patient's Snellen acuity. The user should identify the block of text of smallest size that he can still read, thus setting his near vision acuity. [14], [37], [38]

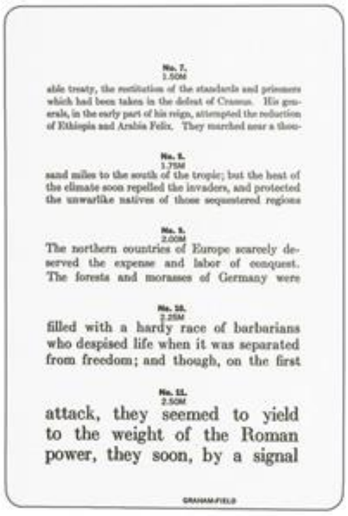

Figure 6. Jaeger Eye Chart

However the testing distance (normally defined as reading distance) and the text size were not standardized when it was first released. Several researchers have since then been developing this work, leading to a better standardization of the evaluation process of reading tests [14], [39] [40], [41]

2.1.1.6 Amsler grid

Amsler grid is a grid of horizontal and vertical lines used to monitor the user's central visual field and detect signs of macular degeneration[42], [43]. The user is asked to focus on the center point of the grid and instructed to report if the grid does not appear uniform.

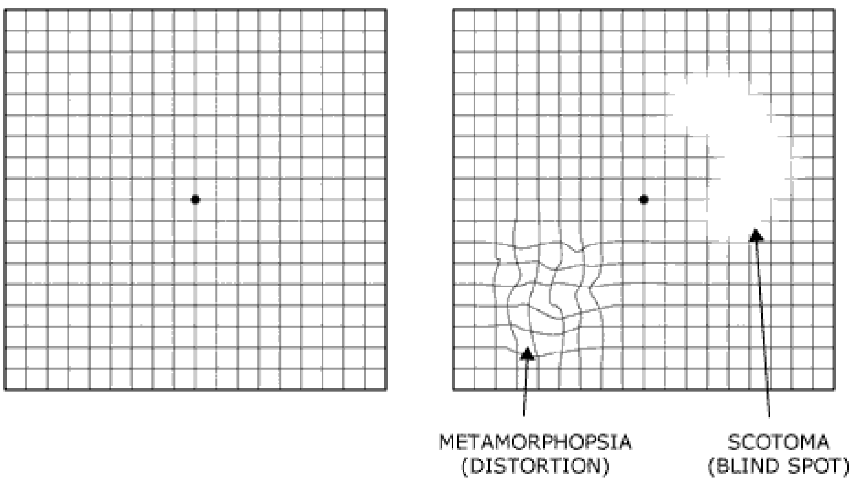

Figure 7. Amsler Grid

Users with normal eyesight will see the undistorted Amsler grid that is shown in the left side of Figure 7. However, if eyesight is problematic he might see some kind of deformation similar to one of those present in the right side (of Figure 7). Metamorphopsia is a type of distorted vision, in which straight lines appears wavy or distorted.[44] Scotoma, also known has a blind spot, represents a defect in the vision field, when an area has diminished visual acuity, but its surroundings have good or completely normal vision. [45], [46]

However, as it has been reported, [4], [43], [47] this test suffers from a number of flaws, including poor validity and lack of ability by the patients to report what problems they actually identify.

2.1.1.7 Ishihara Color Test

The Ishihara Color test [48], [49] attempts to evaluate the possibility of red-green color blindness of a person. It features a number of colored plates, where people with different degrees of colorblindness will obtain different results from people with normal vision. This test is rather simple to take: a circle of apparently randomized dots (with different colors) is presented, for people with normal vision, numbers will be clearly present, while others will either see a different number or, in some cases, nothing at all.

In Figure 8, one of the plates is shown as an example. For people with normal vision, the number 74 Figure 8will be present. However, people with some degree of color blindness might see it otherwise: some may be able to see a 21, while others may actually see nothing but meaningless dots.

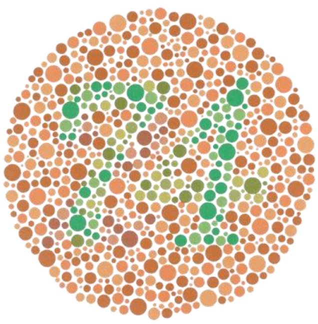

Figure 8. Ishihara Plate 9

2.1.1.8 Vision Summary

This section presented some of the most common vision tests deemed relevant for this dissertation's objective of senses evaluation, including:

- Visual Acuity Charts (Snellen and LogMAR)
- Pelli-Robson Contrast Sensitivity Chart
- Freiburg Vision Test
- Jaeger Eye Chart
- Amsler Grid
- Ishihara Color Test

As mentioned before, these tests were selected due to their role in the clinical vision assessment. Their implementation into a smartphone, using a game-like setting, could prove to be helpful in keeping constant surveillance on vision quality. Likewise, it is believed that all the tests above are an important part of vision evaluation and, as such, they were considered for the purpose of this research.

2.1.2 Audition evaluation

This section focuses on documenting some of the important tests that are currently conducted in the field of audition evaluation. These tests were all considered when deciding which ones would be implemented and given a game-like setting.

2.1.2.1 Pure Tone Audiometry Test

A Pure Tone Audiometry (PTA) test is one of the standard hearing tests used for detecting hearing loss. It is normally conducted in a soundproof booth and with the help of a technician. It is usually conducted for both hears individually, by emitting pure tones through earphones into each ear of the patient and getting feedback whether he was able to detect it or not. By registering the frequencies and intensities the patient can hear, one can detect whether a problem exists and the degree of the hearing loss at that frequency [50].

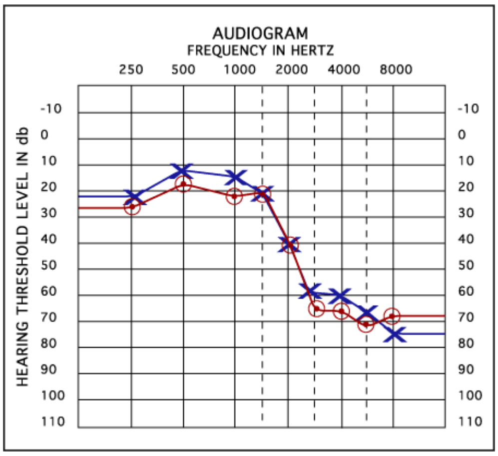

Figure 9. Audiogram.

Figure 9 represents an audiogram obtained from a PTA test. The left ear is represented by the blue line with crosses (X), and the right ear is represented by the red line with circles (O). From the result obtained, one can see that this patient has troublesome hearing, especially at frequencies higher than 2000Hz, since this value falls below the normal level of 20 Db threshold.

2.1.2.2 Speech Discrimination

The Speech Discrimination Test, also known as Word Discrimination Test, is used to assess the patient's ability to not only detect, but also understand sounds (in this case, words). The user is presented with a word, which may or may not be surrounded by a carrier phrase, and is tasked with identifying it. Examples of carrier phrases are: "Say the word" or "Repeat the word". It has both been said that this helps to obtain more accurate test results [51], [52] and that it has little or no effect on the outcome of the test. [53] The test is conducted for both ears together and should be provided in a quiet environment.

The stimulus might be a digital recording or it might be the Doctor simply reading the word. The test is normally conducted at the patient's Most Comfortable Level (MCL) or higher, to avoid lower results due to the inability to hear lower sounds. [54], [55] Prior to conducting the Speech Discrimination tests, the professional determines the Most Comfortable hearing level by providing the patient with a sound and adjusting the sound volume depending on the feedback obtained.

The score obtained on this test is represented in the form of the percentage of correctly repeated words.

2.1.2.3 Hearing in Noise

The Hearing in Noise Test (HINT) evaluates a patient ability to understand speech in noise. The test is conducted for both ears together, the user is presented with speech surrounded by noise, and it is recorded if the user can repeat the stimuli he just heard. At the end of the test, a signal-to-noise ratio (SNR) is generated for each test condition. A signal-to-noise ratio equals how loud the sentences needed to be turned up above the noise floor so that the patient could repeat them correctly in 50% of the occurrences. [56]

2.2 Vision and Audition evaluation in Smartphones

The idea of using smartphones as a medical assistance tool is not new. There are numerous reports of attempting to use these devices, which can be easily carried and used, to support the medical procedures usually conducted. Some of these are meant to be used by medical staff and help them to conduct tests on patients; others are meant to be used by persons wishing to test themselves, even if in most cases these latter tools are not validated clinically.

One study [57] divides the tools available in five categories:

1. Testing Tools, which provide tests that can be performed on one’s self, such as, near vision cards and Amsler grids;
2. Patient education Tools, which is comprised of resources meant to provide the patient with useful eye related information, such as, brochures and videos about medical conditions;
3. Physician reference Tools, which provide the tester with, among others, reference sheets for expected values and grading systems;
4. Physician education Tools, which intends to help physicians to expand or verify their knowledge of common areas, might be important diagnoses to keep on

constant alert for, or questionnaires to assess the presence of commonly encountered diseases;

5. Calculator and other office-based Tools, which is comprised of various conditions calculators, for example, glaucoma risk calculator and back vertex calculator.

References and Calculators might prove themselves useful to work in the background as support for the feedback given to the tested user.

Applications consisting of an audiometry test or vision test can be easily found on the market and, while most require clinical validation, are seen as a way to facilitate patient treatment and communication, as well as data gathering and treatment. [58]

Examples of further investigation which aim at similar objectives will now be presented.

Visual Acuity Testing on the iPad

The work presented in [59] recognizes the advantages of using the iPad for visual acuity measurement, due to its large screen and portability, and attempts to evaluate the results obtained from measuring visual acuity using this device. It performed tests on 85 subjects, measuring visual acuity using an iPad, a Bailey-Lovie chart and an ETDRS letter chart. Upon comparing the results obtained, it concluded that, with measures taken to decrease glare, the evaluation using the tool yielded similar results to those obtained from commonly used charts.

Another study, entitled PlayWithEyes, attempted to evaluate a platform for visual acuity assessment in very young children. [22] It is focused on young children in order to detect early signs of vision problems, seeing as some conditions, if detected early, can be subject to treatment and be reversed. However, testing on young children is not an easy task due to the difficulties of maintaining their attention and cooperation throughout the test. The work led to the creation of a serious game (PlayWithEyes), which provides the child with an iPod he can interact with, and the tester with an iPad which works as the server and allows test configuration. An external display presents Lea symbols, which the child identifies in his iPod (as seen on Figure 10, taken from the published paper [22]). After performing tests in about 200 children, this research concluded that a longer cooperation was achieved with this testing procedure, due to the fun factor the game presented.

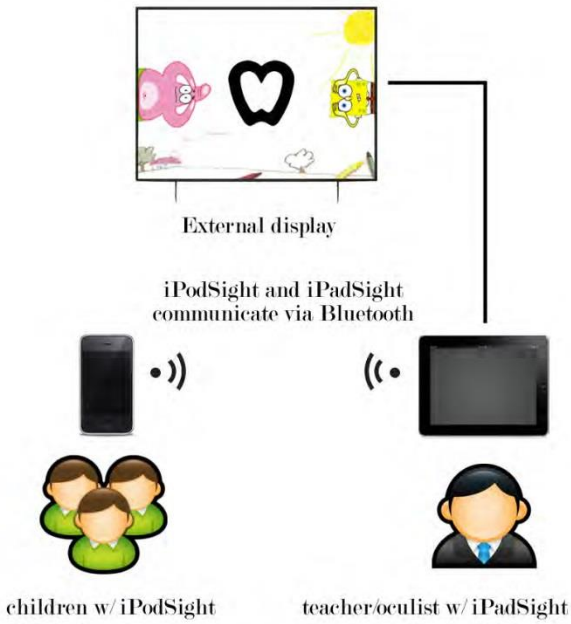

Figure 10. PlayWithEye system

Shape Discrimination

This work is a successful attempt to make shape discrimination test through iPhones, [4] using three shapes. One is different and the user has to identify it. This work is based on previous research conducted in age related shape discrimination, [42], [60] conducted using a computerized version with a similar test procedure, which found that patients with higher degree of age-related macular degeneration obtained worse results in the shape discriminations trials. The handheld version obtained a good correlation of results when compared to its previous version and allowed the user to perform the test by himself. A trial from the test can be seen on Figure 11, taken from the article which documented the study. [4]

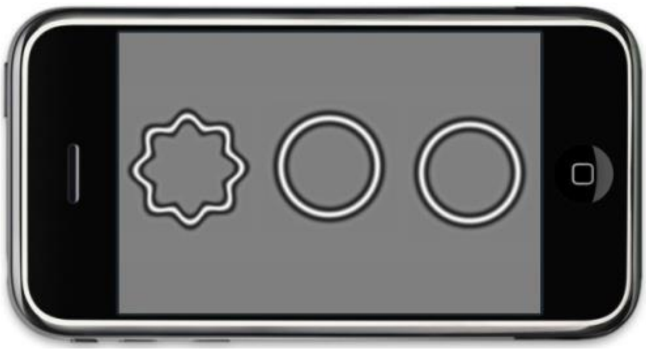

Figure 11. Shape Discrimination trial

Peek Vision

Peek Vision is a system developed for Android phones designed to be a lightweight tool that health workers can use on patients in challenging situations (ex: mobility reduced user, third world countries, etc.). Due to its portability it allows a technician to go to where the patient is, instead of waiting for the patient to seek help. It allows the technician to conduct a visual acuity and visual field tests and, with the help of low-cost hardware, obtain images of the back of the eye (as shown in Figure 12, taken from Peek Vision's website), which are then sent to specialized personnel for evaluation. [61]

Figure 12. Peek Vision in use

One research with a similar component, attempted to evaluate a technique of fundus photography in human and rabbit eyes using a smartphone. [62] By using smartphones' high definition camera and light emission, with the help of already available and portable medical tools, the study was able to deliver a simple technique of fundus photography. It provides a portable tool that can help patients' diagnosis from a distance.

Freiburg Vision Test

The already mentioned Freiburg Visual Acuity and Contrast Sensitivity test [12], [18], [19] is a computer based vision test which, in its first iteration in 1995, targeted only visual acuity evaluation. However it has been constantly developed and was made available in multiple versions, allowing contrast sensitivity testing through fading Landolt Cs or contrast gratings, as well as visual acuity through the use of Tumbling Es or Landolt Cs. The most recent version is available online. Figure 13 shows a contrast grating trial created by the Freiburg Vision Test; the objective of the user is to identify the orientation of the gratings (vertical in the presented case).

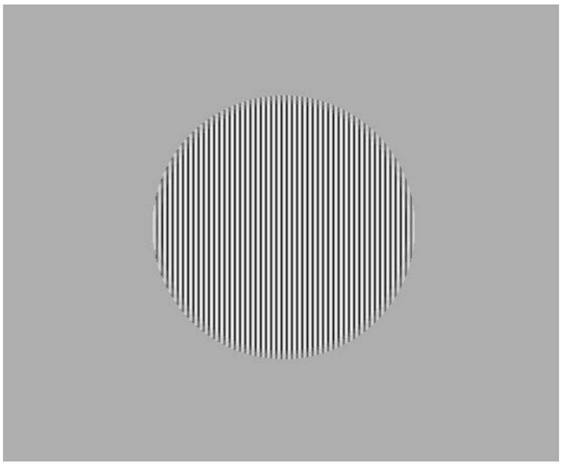

Figure 13. Contrast grating trial

Audiometry Tests

The audiometry test is of major importance for hearing evaluation and is the one which is most frequently performed on patients. Thus, there have been attempts to evaluate its implementation on a smartphone device, as it is documented in the following paragraphs.

One study, which obtained encouraging results, attempted to reproduce a pure tone audiometry on a smartphone and assess its validity on non-soundproof environments. Testing was conducted on twenty test subjects in both a soundproof booth and a quiet room. The results presented a good correlation and permitted to conclude that a soundproof booth wasn't always necessary; [63]

Another test which focused on audiometry testing aimed at developing a similar self-assessment test. This test replaced the traditional use of pure tones by four Korean phonemes which were selected due to their similarity to English phonemes. Testing was conducted in 15 participants and concluded that the results obtained in the phoneme-based test yielded hearing thresholds similar to those obtained with traditional audiograms. [64]

The set of works presented in this chapter seem to indicate a solid connection between senses testing (vision or audition) and the advantages of using portable technology to do so. Many of the research and tools developed aim at helping the professionals to perform their tasks, or allow the users to collect complex data, which is then sent for examination by a professional. Others attempt to provide the user with the possibility of evaluating their own vision or audition. This also shows a solid background between senses evaluation and mobile devices, implying that implementing these features in an Android device, and afterwards giving them a serious game context, can be advantageous to the continuous evaluation of the mentioned senses. While some tests are already present on the iPhone or Android, they lack some of the advantages mentioned before, such as: easy to use and engaging experience for senior citizens, low-cost and self-contained system and the absence of external help requirement.

2.3 Challenges

The fact that we are attempting to reproduce these tests in a much different environment brings forth a number of challenges that must be considered.

- **Surrounding Environment** – The environment the user is in might influence the outcome of the exams. The environment can influence both vision exams (for example, by light reflecting on the screen) and audition exams (for example, due to noisy situations that require a higher volume than usual). Therefore, it is important to try to mitigate this impact as much as possible by putting in place features that either reduce it or give alert messages to the user so that he can avoid performing the test in those conditions.
- **Hardware restrictions** – The exams developed have to be reproducible in the user's smartphone. While powerful, these devices have obvious limitations: small and variable sized screen when compared to a standardized vision chart, as well as sound volume output limits that are relative to each device. This is a factor that weighs heavily in the design process of the methodologies used, influencing exam and game design, as well as influencing the need of additional features such as a proper calibration.
- **Additional hardware –** one of the goals is the development of an easy to use and easily available system. As such, the exams implemented need to be as self-contained as possible. Any additional hardware, beyond the smartphone, is considered an extra restriction to the application usage. Requirement of extra hardware generates different problems and considerations:
    - System cost – the component must be cheap and easily obtainable;
    - Complexity – the component must be easy to operate, in order not to increase the complexity of the system, reduce satisfaction or affect test performance. All of these could lead to inaccurate exam results, obviously something that should be avoided.
    - Testing environment problems – by adding an extra component that could perhaps restrict the setting on where the user can perform the exam, we are possibly diminishing the frequency of the test. Ideally the user should have all the tools available every time, so that he is always able to perform the exams.

All of these are factors that might influence the user experience or exam results and as such, must be taken into consideration during the design and development process.

2.4 Technologies

As requested by Fraunhofer's proposal, the prototype developed is a smartphone application for Google's mobile operative system, Android. The target devices are common mid-level devices; as such the system should be able to run on these devices without requiring high-end devices features to run flawlessly.

In order to develop these prototypes, the official Android SDK provided by Google Inc. is used. The research conducted and the envisioned exams revealed the need to emit sounds at certain frequencies. In order to generate the sound and then play it to the user, Android's AudioTrack API is used. To make the exams more precise, the user should wear a headphone set, therefore reducing ambient noise, increasing precision and allowing for a left-right ear differentiation of sound emission. Also related to the hearing exams, namely the PTA, an audiogram-like graph will be needed to be drawn. For graph drawing two open source projects were analyzed: the GraphView and AChartEngine library. After some testing with each library, the AChartEngine was the one selected due to its better ability to produce the desired graphs and ease of use.

To continue taking advantage of Android's APIs, the Graphics API is used for the shapes creation and animations, not only in the shape discrimination exam, but also in the gamification process of the remaining implemented exams.

Considering the challenges expressed previously, a number of features are required to help in the testing process (these are documented in more detail in chapter 3.6-Other Features):

- Distance measurement. To implement this, the open-source library OpenCV is used. Due to its image processing capabilities it enables us to capture an image through the smartphone's camera of the user's eyes, which, after calibration, is used to get an estimation of the user's distance to the device.
- Environment light adequacy. The phone's light sensor is used, so that the user can get feedback on whether the current environment might have an impact on the exams. If it does, it might influence exam reliability and the user will be advised to avoid performing the exam in those conditions.
- Screen brightness. During vision tests, the device brightness setting is overridden and set to the maximum possible value. This assures that the same brightness setting is always used and provides the user with a better visual experience.

2.5 Summary

In this chapter the research that was conducted and deemed relevant is documented, as well as the findings. Both audition and vision were analyzed and a set of tests was presented and explained. This explanation was meant to give an idea of the kind of tests that are worthy of being reproduced. Similar works were presented, that can be useful as a validation of techniques used, or to present some ideas that can be worked on and improved.

Taking into account Fraunhofer's needs for the developed prototypes and considering the tests identified in the research step, a brief description of the technologies used was also given.

A more detailed explanation of how the research conducted was used for this dissertation development will be presented in the next chapter.

Table 1 is a brief overview of the tests examined in this chapter, their objective and special considerations that need to be taken into account during their implementation.

Table 1. Overview of Tests

Test Name	Evaluation Target	Special Considerations
Snellen Chart	Visual Acuity	Patient-Chart distance (approximately 6 meters)
LogMAR Chart	Visual Acuity	Patient-Chart distance (approximately 4 meters)
Pelli-Robson Chart	Contrast Sensitivity	Patient-Chart distance (approximately 1 meter)
Freiburg Vision Test	Visual Acuity and Contrast Sensitivity	Minimum letter size
Jaeger Eye Chart	Near Vision (Reading)	Minimum letter size
Amsler Grid	Central Visual Field	None
Ishihara Test	Color Blindness	None
Audiogram	Hearing Loss	Left/Right ear differentiation
Speech Discrimination	Speech understanding	Prior calibration of Most Comfortable Level
Hearing in Noise	Speech in Noise understanding	Sound vs Noise relation

Chapter 3

Senses Evaluation: a Methodology

This chapter presents and details the methodologies proposed to evaluate the users' vision and audition. First it will focus on detailing which tests were chosen, why they were incorporated or discarded, as well as the main decisions that were made regarding their implementation. Afterwards, it will document each test development process and show the final version of the games which ultimately led to the present version of the prototype.

3.1 Overview

In order to develop a senses evaluation system that is suitable for senior citizens, a research on the available tests was first conducted. A number of tests was found, some of them with similar goals. After considering the implications of each one and taking into account the timespan provided, a selection was made for the ones that would be implemented.

Due to the target users of the system (senior citizens), special considerations were warranted. [65], [66] This had an impact on the planning and development of the games and general flow of the prototype, in an attempt to provide the user with a comfortable experience that would not limit their ability or desire to use the application on a regular basis.

Before a game was created, the features that would be required were developed and later it would then be given a more game-like approach by enhancing the user interaction and by making the overall aspect of the game more appealing. The selection and development progress of each sense exam is documented in this chapter.

3.2 Sense Tests Selection Considerations

The selection process of which tests should be chosen was made based on a number of factors that need to be considered. Those factors include:

- Device and the additional hardware restrictions mentioned in 2.3-Challenges (page 19).
- Relevance of the test – which was based on the overall use of the test on the current clinical settings. The tests that are more common received a higher degree of relevance over the tests that are only used in very specific scenarios. It is important to notice that, even though these tests might be the most used, it does not mean that one singular person is regularly tested at them. If the test is deemed so important that it is frequently administered to the general population, then the system should attempt to track the progress of the user at that test, in hopes of quickly detecting changes.
- Time commitment – understandably, given the relatively short time span that was available to develop this type of work, time available played a big part on the selection of tests. The time required to go through all the steps of a test was considered and compared against the other tests and their respective advantages. Some tests had to be cut due to the already envisioned allocated time and the time commitment required for that test.

The following section will highlight what tests were chosen to be developed, their characteristics and other considerations.

3.3 Test Selection

This section presents the selection process of the tests that were gathered during the research step seen in the previous chapter.

3.3.1 Distance Visual Acuity

It was decided that a exam intended to measure distance visual acuity would be developed. Its impact on everyday life is well known [21], [67], [68] (ability to see objects at a distance, loss of independence, etc.). These reasons made this exam one of the crucial vision characteristics to analyze. Like mentioned in Chapter 2, there are different tests available, but most of them share a similar procedure (for the chart based tests) and considerations, with some slight variation such as: optotypes used, size reduction steps, character spacing and number of characters per line. However, it also presents various problems:

- A smartphone device is much smaller than any distance visual acuity chart normally used to conduct this test.
- The usual testing distance is much larger than that which the user can achieve by simply holding the device in his hand and distancing it from himself.

- The optotype size is restricted both in maximum size (by the physical screen size) and minimum size (by the smartphone resolution, at which point the optotype will be deformed).

These considerations were taken into account and a testing procedure was elaborated that attempts to mitigate these problems. This process is documented in Section 3.5.4.

3.3.2 Near Vision

Near vision is another important part of one's vision. While a person might not be able to see clearly at long distance, they may or may not have problems with objects at shorter distances. For near vision several tests are available such as: reading chart (like the Jaeger eye chart) and near visual acuity charts (similar to a Distance Visual Acuity charts, but for use at 40cm). Taking into account the desire to provide the user with engaging games, an exam similar to the Jaeger eye chart was chosen to be developed (over another optotype based chart like the near visual acuity one), in an attempt to keep the games from being overly repetitive and allowing the evaluation of a different metric (reading acuity over near visual acuity). So, the exam developed focuses on evaluating the user's ability to read a sentence and it is also important to take into considerations the device size (limiting sentence size), the distance (which is much shorter than the Distance Visual Acuity test's one, therefore easier to replicate) and the letter size. The steps to assure the user's ability to read a given sentence are covered in Section 3.5.2.

3.3.3 Contrast Sensitivity

Contrast Sensitivity tests are not as frequently administered as Visual Acuity ones. However, there are several research works linking low contrast sensitivity scores to the presence of cataracts, even when Visual Acuity scores are normal. [69]–[71] Also, taking into account the work developed by Lord et al. [1]–[3] which suggests a relation between falls and low vision (both acuity and contrast sensitivity), a contrast sensitivity exam was deemed relevant to the project this dissertation is inserted into and was selected to be implemented. Again, testing distance (which can be large, like in the Pelli-Robson test, or short, like in the FrACT) and the environment where the user is in (especially light reflection on the screen) are important factors here.

3.3.4 Shape Discrimination

Taking into consideration the work developed by Wang et al. [4], [42], [47] in which they created an iPod/iPhone based application, which provides the user with a shape discrimination test, in an attempt to evaluate the user's vision and find early traces of early stage diabetic retinopathy, before it can evolve, a decision was made to develop an Android based application with a similar procedure. Once again it is important to consider the shape size and the testing distance (similar to that of the near vision exam).

3.3.5 Pure Tone Audiometry

Much like the Visual Acuity test is to vision assessment, the PTA test is one of the key tests for audition evaluations. It is regularly administered to patients in an attempt to evaluate their hearing performance and perform a diagnosis of potential problems. This is why, the PTA exam was deemed too relevant to be left out of our testing options. By continuously tracking his hearing ability, the user may be able to detect potentially arising problems and be alerted to the need to take an action as soon as possible.

3.3.6 Speech Discrimination

This exam was selected due to its importance in day to day life of a person. While the PTA score may lead users to believe they can hear sounds well, that doesn't mean they can identify what the sounds actually mean (in the case of speech). This exam attempts to help to prevent such case, by providing the user with a mean to test his ability to identify words at a comfortable volume.

3.3.7 Non-implemented Tests

In an attempt to keep our testing range manageable, following the guidelines set in section 3.2, the following tests were discarded:

- Amsler Grid – This test was not implemented since it is easily obtained and easily usable by the user in its traditional form. Due to the fact that it is a simple and relatively small paper, which doesn't require any variability to remain valid, it wouldn't have as many advantages in being converted into a mobile version, when compared to the tests that were indeed selected.
- Color blindness – Similarly to the Amsler Grid, this test is easily available to the user and doesn't require any external help. As such, there was a less pressing need for a mobile version of this test.
- Speech in Noise Understanding – Due to the fact that the PTA and the Speech Discrimination test were already selected to target hearing evaluation, the decision was made to leave the Speech in Noise understanding out of the test range.

3.4 Considerations

The following are some considerations taken into account when designing the methodologies used.

3.4.1 Quick Testing

The aim of the system is to allow the user to continuously monitor his vision and audition. However this does not mean the user must perform every single exam every day. Such a task would be cumbersome to the user and could potentially lead to two major problems:

- A more disengaged and careless use of the exams, resulting on exam results with lower scores. Larger scores will be achieved if the user is trying his hardest and is not tired of a continuous array of testing;
- An eventual reluctance to use the application, due to the time commitment required to fulfill all its requirements.

A choice was then made that each exam would be developed as a single game, instead of a combination of different exams at the same time, giving the user the choice of which one to play at any time. This allows the user to keep a healthy balance of his desired exams, performing them in a frequent basis, even if alternating between games at each test run.

This provides the user with quick games, which can be performed at any time, therefore hopefully increasing the exams frequency instead of limiting it to long time spans of free time. Additionally, this allows the user to be interested in monitoring one sense (or certain characteristic of one or more senses), and so easily focus on performing only the games which target that sense.

3.4.2 Ease of Use

Due to the characteristics of the target audience of the system, which are different from the more regular users of mobile devices, one must be careful about the level of skills required to use the system. [4], [65], [66], [72] Because of this, the system was designed with the aim of providing a simple interaction experience. In each screen the user is provided with few options, in an attempt to not overwhelm him with information. Every interaction expects nothing more than clicks or presses on large buttons, removing small sized elements, as well as drags or slide movements that could present problems to users with limited movement capabilities.

3.4.3 Exam Procedure Validity

In order to obtain accurate exam results two types of calibration have to be considered: the PTA test and the vision tests.

3.4.3.1 Calibrations

Both these exams are very susceptible to device and environment related influence, as such they warrant special considerations:

- The **Pure Tone Audiometry** exam is especially focused on finding the minimum volume a user can hear. However, without prior calibration, it is extremely hard to know what volume the user is actually being exposed to. Android is largely available in an enormous quantity of devices and each device will produce different volume outputs. Then, one still has to consider the fact that, in order to allow for ear differentiation and noise reduction, the user must wear an earphone/headphone set. However, this raises two other problems:

• Each set will vary in its own volume output;

• Earphone and headphone will have different impacts. The earphone placement is vastly more variable than headphones. If the user opts for earphones over headphones, the user will be losing the noise reduction (because headphones cover the ear in a better way) and the constant placing that headphones normally offer (it is much easier to place headphones in the same way each test run, than it is for earphones).

In order to try to solve these problems the use of headphones is recommended and until a calibration can be made to the system, it will use a device relative volume reference. The minimum volume required for a frequency will be represented as a volume index for a certain device-headphone set combo. While this will not allow the user to determine the real volume he can hear, it will allow him to track his relative performance (assuming he follows the guidelines provided) and identify potential flaws.

- The **vision exams** will produce different results depending on the screen specifications and the lighting the user is faced against. The device's screen size and resolution will impact maximum and minimum letter size, as well as influencing the contrast differentiation possible. Different screens will also react differently to light reflection, especially sun light. In order to diminish the impact this has on exam results, three features were developed:

- Size calibrations. The user is able to calibrate the maximum and minimum size of letter presentation, with the help of the system.
- Light Sensor. This is used to try to assess the current illumination conditions and advise against performing the exam if they are considered adverse.
- Brightness. For vision exams, the brightness of the display is always automatically set to maximum, overriding the previous setting of the device. Thus improving view conditions on most situations, therefore reducing variability. After completing the exam, the device's brightness is returned to the original value.

Additionally, vision exams also depend on the user distance from the device. In order to help the user in maintaining the proper distance, OpenCV is used to estimate the user-device distance and give advices on if he should move closer or farther away.

3.5 From Tests to Games

In this section the development process of each exam will be documented. When applicable the first iterations of the exam will also be presented, showing the development of the exam from the simple procedure testing phase to the final game form.

3.5.1 Shape Discrimination Game

This game is based on the work of Wang et al. [47], [60] and Wilkinson et al. [73] and, as such the first iteration of it was a re-implementation of Wang's work on the iPod/iPhone to an Android setting. [4], [47] His work focuses on the discrimination of shapes that can take different forms. However, a choice was made to use the circular 4th derivative of Gaussian contour (CD4) patterns and their deformations [73], [74] since it most closely resembles the work he performed on handheld devices. The CD4 patterns can be generated using the following formula [60], [73]:

$$D4(r) = C * (1 - 4 * \frac{r - r0^2}{\sigma} + 4 * \frac{r - r0^4}{\sigma}) * e^{-\left(\frac{r - r0^2}{\sigma}\right)} \qquad (2)$$

Where *C* is the pattern Contrast, σ determines peak spatial frequency, *r0* is the mean radius and r is representative of the current position being evaluated.

When a distorted shape needed to be created a radial deformation would be applied to its radius *r0* in Formula 2, using the following formula:

$$R = r0 * (1 + A + \sin(rf * \arctan\left(\frac{y}{x}\right) + \emptyset)) \qquad (3)$$

Where *A* represents radial amplitude, *rf* represents radial frequency, *y* and *x* are the current position being analyzed and ∅ is the phase of modulation.

Following the work developed by Wang on the iPod/iPhone the pattern size is calculated so that it makes a visual angle of 2 degree at the testing distance. In order to keep a constant distance among exams which require the user to hold the device in his hand, the distance of 40 centimeters was selected. This distance was chosen because it is the distance usually considered for the Near Vision test (as reported in the distance selection section of 3.5.2) and it does not greatly differ from Wang's chosen distance of "at an arm's length" (which he defines as "approximately 500 mm"). [4]

In Figure 14, a first draft of the intended use can be examined. In this game, the user will be tasked with finding the irregular shape. Upon selecting the correct shape two times in a row (to avoid random guessing) the user would be shown to a new trial with a less deformed shape. If he got it wrong, the shape would be more deformed in the next trial (thus easier to spot). The deformity step (the amount of deformity reduction or increase) would be higher at the beginning but at each reversal the step would decrease, getting closer to the user threshold. A reversal is defined as an error when the last action was a deformity increase (or unchanged – meaning the user last actions were correct responses) or two correct answers after a deformity decrease (meaning the previous action was an incorrect response).

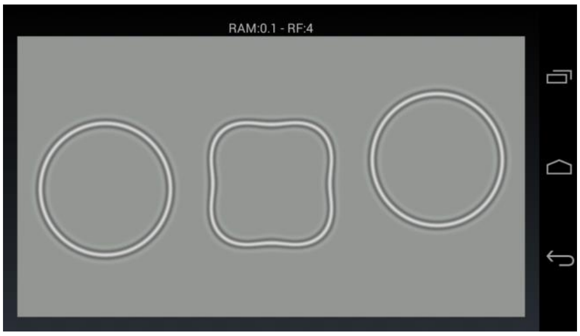

Figure 14. Shape Discrimination (First Mode)

In the next iteration, to give the user a more entertaining game, a different approach was taken. The user was shown six shapes; these shapes have a Radial Amplitude (RAM) deformity between 0.1 (most deformed) and 0 (undeformed CD4 pattern) with a constant deformity reduction between them. This means six shapes were using a RAM defined by the following formula:

$$\mathrm{RAM(i)} = \left(\frac{\mathrm{RAM0}}{(\mathrm{n}-1)}\right) * (\mathrm{i}-1) \quad (4)$$

Where *n* is the total number of shapes, *RAM0* is the maximum Ram, *i* is the index of the shape being examined and $\mathrm{RAM(i)}$ is the resulting shape's RAM.

These shapes would then have their order randomized and the user will be tasked with ordering the shapes from the most deformed to the least deformed (circle). After each successful trial, the *RAM0* value would be decreased and a new trial will be presented.

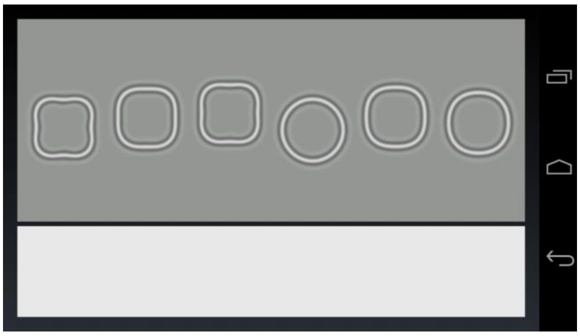

Figure 15. Shape Discrimination (Second Mode)

However, this provided the user with a considerable challenge immediately in his first attempt, because the second least deformed and the least deformed shapes are already hard to distinguish. Additionally, the free space available would be very small with this procedure. Despite the fact that we can see the six shapes on the screen in Figure 15, the shapes are not in their recommended dimensions (meaning that when they are, there wouldn't be enough space to prevent overlapping). The initial RAM could have been set to a higher value, but while examining the options, a way to provide the user with a more familiar testing environment was found. Based on the traditional memory game where the user is tasked with matching identical images, a game was created. By providing the user with six shapes (all visible unlike the memory game) divided in three pairs of two equally deformed shapes, the user starts the game with easier shapes in a simpler testing procedure. Three shapes are created with Formula 4 (but now with $n = 3$), then those shapes are cloned, totaling 6 shapes and they would have their order randomized.

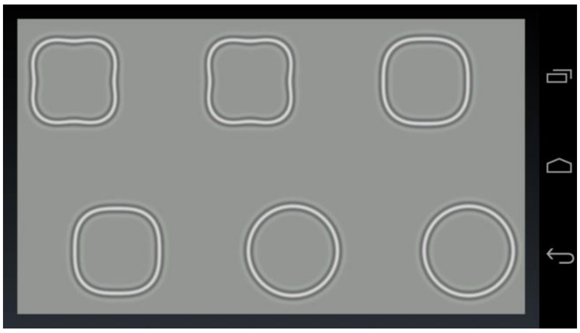

Figure 16. Shape Discrimination (Final Mode)

Looking at the shapes the user's objective is to make pairs of similar shapes, starting by matching the most deformed and ending in the least deformed (the circles). It is visible that instead of placing all the shapes in a single row, in this exam two rows were created, allowing for bigger shape sizes (in this image, shape size is accurate). This would be the final iteration of this game.

These shapes are drawn on the canvas in runtime, the resulting image would be stored so that after the initial drawing of the shape with a certain RAM, one could reuse it without repeating the calculations required, thus reducing the trial generation time. The answer detection is performed by a touch listener. Upon selecting a shape, a square is drawn around it and, upon selecting the correct match, both of them are made invisible to the user.

3.5.2 Near Vision Acuity

For the Near Vision exams, like mentioned before, a reading text was the chosen evaluation method to be implemented, focusing on assessing the user's ability to read sentences at certain sizes. [75] This was chosen instead of another optotype identification method (which was already being developed for the Distance Visual Acuity), mainly because it produces completely distinct games and experiences, hoping that this does a better job at motivating the user than two similar games would. By doing this, a different testing procedure and different metric (reading ability over near visual acuity) is obtained from the user. The letter size notation used is the “M-units” notation, a term coined by Louise Sloan, where 1 M-units subtends a visual angle of 5 min of arc at 1 meter, [14], [76] as such n M-units will subtend 5 min of arc at n meters. Much like the card versions of this test, it is intended to be used at 40 centimeters from the user's eye, thus being easily performed by holding the device in the user's hand. Upon finding the user's minimum readable size, one can estimate his visual acuity by using the following formula [77], [78]:

$$\mathrm{VA} = \frac{d}{\mathrm{M}} \qquad (5)$$

Where VA is visual acuity score, d is the distance from the user to chart or device in meters, and M is the text size of the smallest letter size the user is able to read in M-units.

This exam went through different versions. In the first version, it faithfully emulated the traditional test, even though it was a very simple procedure. The user was presented with a sentence and he either gave feedback to the system that he could read the sentence, or not. If he could read it, a smaller one would be given, until his minimum readable size was found or the devices resolution would not allow it to continue any further. If the later happened, this fact would be recorded to indicate that the user reached the device's ability and not that the value achieve was the user's true limit. An image of this earlier version of the exam can be seen on Figure 17.

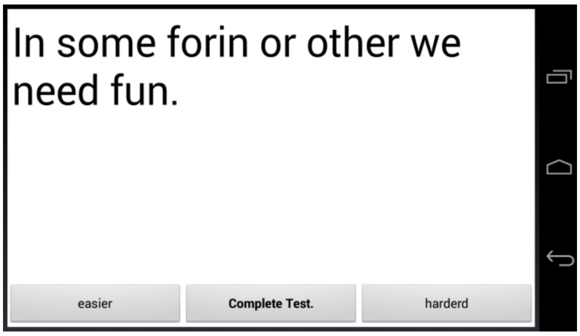

Figure 17. Reading Exam (Initial Mode)

However, this procedure wasn't seen as a very engaging activity, leaving the user with only the task of answering “yes” or “no” to the implied question of “Can you read this?”. Additionally, this would make it very easy for the user to feed incorrect information to the application by simply indicating he is able to read it, even though he might not. So, in order to attempt to fix these two points, a more game like experience was created. It was clear that proof that the user could read the sentence would be an important step. So, two options were analyzed, either: the user would read the sentence out loud, it would be analyzed and checked for a positive match, or the sentence would have missing letters and the user would be tasked with completing it. The advantages and disadvantages of both were waged against each other and the later ended up being the one chosen, since it was believed any form of speech impediment or accent could lead to inaccuracy in the speech recognition process. The sentences used in this exam should be simple enough as not to provoke confusion on the user. Actually, the purpose of the exam is to evaluate his ability to see the text clearly enough to read it and not to evaluate his knowledge and performance as a reader. The user was given a sentence and, in normal reading order, should choose the letter (by clicking on it) that he deemed the correct one in the missing spot. An example trial of this game is show below (Figure 18):

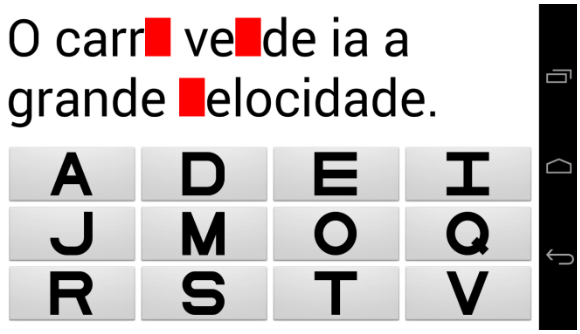

Figure 18. Reading Exam (Final Mode)

Since sometimes more than one letter would fit a given spot, the user was allowed to make up to three tries on any giving missing letter. If he however got 3 incorrect tries in a given spot or 5 incorrect tries in one sentence his exam would be over and his minimum readable size would be set to the last successful sentence. This margin for error also helped protect the user against accidental clicks. This would be the final version of this game.

The sentences selected for this exam followed some guidelines:

- They have to be small, sharing a similar sentence size;
- They must contain simple words and a straightforward meaning, in order to be easily understood;
- There must be three missing letters per sentence, but no word would have more than one missing letter;

Additionally, prior to performing this exam, a calibration should be made to the device. In this calibration the user is tasked with specifying two things:

- Maximum possible text size. In the calibration window all allowed text sizes are shown. If the text is too big and overlaps with the buttons the user should instruct the application of this fact. However, this step is only a way to speed up the game initiation process; actually, a mechanism was implemented to, when initializing the Reading Game, check for text collision with the buttons, reducing the letter size to the next level if detected.
- Minimum text size to be used. In the calibration window the user is tasked with finding the minimum text size to test. When the minimum size before distortion was detected, the user

is informed of this fact. However, he is given the choice to allow that size if he thinks its distortion is too small to affect reading performance.

Upon performing the exam, this calibration is saved on the record of the exam result. The meaning is that if the user reached one of the limits, it will be known that a worse or better score was not possible, and cannot be confused with the user's definite reading ability. The allowed sizes (in M-units) prior to calibration are: 12.5, 10.0, 8.0, 6.3, 5.0f, 4.0, 3.2, 2.5, 2.0, 1.6, 1.25, 1.0, 0.8, 0.63 and 0.5. These were selected after analyzing multiple reading cards, as well as near vision acuity cards, and because they followed a uniform progression in letter size. [14], [77], [78]

For the exam generation, very simple elements from Android's standard APIs are used: Text View and Buttons. To generate the helper red square a SpannableStringBuilder is used, substituting the intended missing characters with bitmap images representing red squares. This is presented on the available Text View. The exam interface was purposely kept simple, with black text on white background, generating maximum contrast, so that reduced contrast does not influence reading ability.

3.5.3 Contrast Sensitivity

For testing contrast sensitivity, the work developed for this exam was based on the Freiburg Visual Acuity Test (FrACT) [12], [19], [79] and the Pelli-Robson Contrast Sensitivity Chart. [20], [35] The user is presented with an optotype which varies in contrast with the background. Similar to the FrACT, the Landolt C was chosen as the exam’s optotype. Using an orientation based optotype helped to reduce the exam complexity since it does not force the user to look through a number of options until he finds the correct match (which would happen if Sloan letters, Lea symbols or similar optotype were used). The Landolt C is ideal to the envisioned exam procedure, even better than the Tumbling E, since it is easily rotated in 8 orientations (whereas the Tumbling E will normally only have 4 options). The way this exam works is: a Landolt C is presented to the user in one of eight possible rotations: up (90º), down (270º), left (180º), right (0º) or the four diagonal orientations between them (45º, 135º, 225º and 315º). The intended testing distance, like in the former exams, is 40 centimeters with the size of the optotype remaining constant in 2.8 degrees of arc. [80] In the beginning the optotype contrasts very highly with the background, being black against a white background; however, as the user successfully identifies its rotation the optotype, its color will be increasing fainter (by increasing its transparency), gaining a more grayish look (while the background remains white), therefore reducing contrast and increasing difficulty.

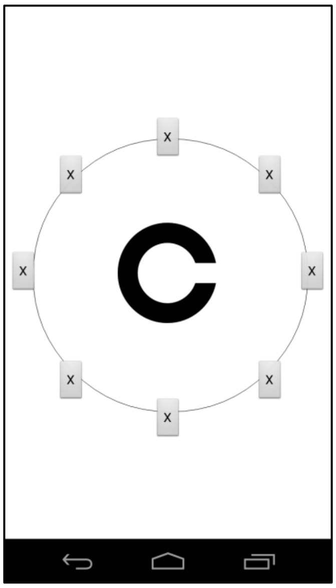

Figure 19. Contrast Exam (Initial Version)

In order to avoid random guessing the user must correctly answer two trials in a row before being given a harder challenge; if he gets one trial wrong he is immediately provided with an easier challenge. The contrast reduction step is high in the beginning but, similarly to the Shape Discrimination procedure mentioned above (Section 3.5.1), at each reversal the step is reduced, therefore getting closer to the user's threshold. Upon reaching 6 reversals the threshold is set, the users Contrast Sensitivity is revealed in the form of Weber Contrast [70], [80] and an assessment is made.

The intended procedure did not change between the gamification process of this exam; however, the appearance was changed to be more entertaining to the user. A background image was drawn as the background. However, in order not to affect contrast, the optotype was surrounded by a fully white circle on top of semitransparent white circle. The buttons were replaced with apples that the Landolt C is trying to “eat”. The user should click on the apple which is in line with the Landolt C's opening. If it is correct, that is, if the Landolt C makes a direct line in that direction and its opening (not his sides) touches the apple, a correct answer is scored. An example of this exam is shown below (Figure 20):

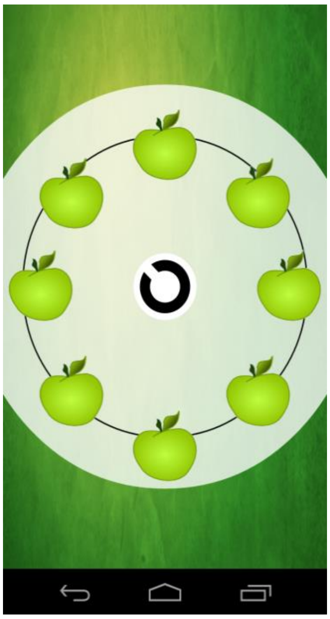

Figure 20. Contrast Exam (Final Version)

The rotation is never the same in two consecutive trials, meaning the user has, at best a 1/49 (1/7 * 1/7, because the user always knows one of the eight options is not possible because it was used in the previous trial) chance of going to the next level by answering randomly to both trials at a given contrast.

For the interface animations, again, a combination of Image Views (for the apples), canvas drawing (for the background and circles) and Text Views (for the optotype used) which are moved by Runnables, were used. The optotype has its contrast reduced by reducing its alpha value. The contrast reduction is made in different steps, depending on the number of reversals (represent as *Rev*), using the following formula:

$$\text{Step} = \frac{40}{2^{Rev}} \quad (6)$$

The resulting *Step* value is rounded down and applied to the current alpha (out of 255), either increasing it or decreasing it depending on the whether the difficulty is being decreased or increased, respectively.

This would be the final version of the game for this prototype.

3.5.4 Distance Visual Acuity

The Distance Visual Acuity exam presented the hardest challenge of the vision exams, due to the natural need of a large distance between the user and the object of the test. The distances used for this kind of test are normally ranged between 3 and 6 meters [5], [6], [81], distances far greater than achievable by holding the device at arm's length.

One option would be the help of another person in order to take the exam; however, this was something that was refused from the beginning of this dissertation: the user should be able to conduct all the exams available without any need for outside assistance.

Another option that was considered was using an auxiliary screen for this exam, through the development of a web application. This application would allow the exam to be presented on a large screened device (computer, tablet or similar) and the smartphone would be used as an input method of the user's answer. However, this option was discarded because it would prove as an extra burden to the user: both economic (through the need to own or acquire such device) and complexity wise (the user would have to be able to use both devices effectively, instead of just one). As such this option was removed.

Borrowing from the techniques used on smaller clinical offices that lack the desired lengths, in which case mirrors are used to double the viewing distance, [11], [12], [14], [38] effectively achieving the expected distance, the same solution was chosen to be incorporated with the to-be-implemented game. The user has to face a mirror and hold the phone with its screen facing the mirror. However, this presented the problem that the user, now not being able to look directly at the screen as usual, would have its interaction with the device taken to a higher level of difficulty. Considerations were made both on the optotype to use:

- Letters
- Symbols
- Direction-based optotypes (Landolt C or Tumbling E)

And on the way the user could input his desired answer:

- Voice recognition;
- Button pressing;
- Device movement based answering;

Speech recognition was discarded for the same reason that was mentioned in Near Visual Acuity. Button pressing presented considerable problems since the buttons would also have their perceived size reduced, due to the test distance. Also, the coordination required to press the correct button, while seeing only it's reflection in the mirror, would be greater than usual. The remaining input option ended up being device movement. The selection of letters could prove to be a slow task using only the smartphone movement (again, made harder by only using the reflection in the mirror), which left us with symbols and direction-based optotypes. From these, a way to mix the options available was envisioned and eventually led to the selection of direction-based optotypes as the final answer, due to their testing purpose and the selected input method.

The selected optotype would be rotated in one of four directions (Up, Down, Left, And Right) and the user would rotate the device in his hand, until it was in the desired position. The desired position was arbitrarily chosen to be the upward position, meaning: the Landolt C would have to have its opening facing the ceiling, and the Tumbling E would have to have its legs pointing to the ceiling. To be easily differentiated from the Contrast Sensitivity Game (which uses the Landolt C), the Tumbling E was selected for this exam. However, the option to change between letters is present, in case of the desire to change the optotype selection.

The choice was made to use only four directions and exclude the intermediary positions due to this being less prone to uncertainty and errors while rotating the phone to the desired position, which also removed one of the advantages of using the Landolt C over the Tumbling E. To allow the user to perform the task of rotating the phone in a controlled way and to avoid undesired answer input by the user, it was defined that the user would need to rotate the phone to the desired position, and then make a long press on the screen. The long press was chosen, because it was seen as a more secure option over the simple click, since it removes the scenario where the user accidentally clicks on the screen while rotating the device.

To start the exam, the user places himself at the desired distance from the mirror, selects the game, selects the eye to test and then flips the screen to face the mirror. From this point on, until the game is over, the user must not look directly at the phone again. The user must cover one of his eyes (so that each eye is tested separately); an eye patch is recommended, but anything that prevents the use of one eye (while not reducing the other eye's seeing ability) will produce the same result. From this point on, the user will be presented with a Tumbling E, rotated in one of the mentioned orientations. He must rotate the device in his hands, until the legs of the E are facing the ceiling, at which situation he must make a long press on the screen, indicating that this is his final answer. Similarly to the other games, if he gets two correct answers in a row, the difficulty is increased; if however the answer was incorrect, he is provided with an easier trial. The rotation is never the same in two consecutive trials, meaning the user has, at best a 1/9 (1/3 * 1/3, because the user always knows one of the four options is not possible because it was used in the previous trial) chance of going to the next level by randomly answering to both trials at a given size. This exam ends in one of three ways:

- The user fails the easiest level (maximum size that fits in the screen) twice, in which case his score is recorded but identified as inconclusive since his vision might be worse than that.

- The user completes the hardest level (minimum size reproducible before distortion) twice, in which case his score is recorded but identified as inconclusive since his vision might be better than that.

• There are 4 reversals. Similar to the previous games, a reversal is defined as a change of test difficulty. This means a reversal is made when a user answers correctly (twice in a row) when he previously answered incorrectly, or when he answers incorrectly when he previously answered correctly (at least twice in a row).

This last point helps is an additional security measure, meaning that even if the user is able to proceed to the next level without being truly able to see the optotypes, he is most likely going to be tested at that same size again (because it will be even harder for him to answer correctly again, which will lead to a decrease in difficulty). Additionally, it also helps to prevent a scenario where an incorrect answer was introduced by mistake.

Figure 21 shows an image of a trial. The user's correct answer would be obtained by the device rotation, as shown in Figure 22:

Figure 21. Distance Visual Acuity Exam (Trial Given)

Figure 22. Distance Visual Acuity Exam (Correct Position)

To reduce the number of unexpected events during the manipulation of the device, the sound buttons are disabled (meaning the Android's default volume bar won't appear) and the back button does not return immediately (default Android's behavior), giving the user a dialog window to choose whether he really wants to quit. The remaining buttons (Home, Power, and Recent) cannot be overridden on non-rotted smartphones, due to Android's security measures. However, the user can in these cases continue the test where he was after pressing one of these buttons, the same could not be said of the Back button's default behavior.

The distances this exam is allowed to be taken at, are 1.5, 2 or 3 meters, which when in front of a mirror reproduce the test at 3, 4 or 6 meters, respectively. The allowed sizes for optotype presentation, prior to calibration (identical procedure to the Near Vision Calibration), are: 37.5, 30.0, 24.0, 19.5, 15.6, 12.5, 10.0, 8.0, 6.3, 5.0, 4.0, 3.2, 2.5, 2.0, 1.6, 1.25, 1.0 and 0.8 M-units. These were selected after analyzing multiple charts and because they followed a uniform progression in letter size. [14], [77], [78]

3.5.5 Pure Tone Audiometry

As a person gets older, their hearing performance slowly degrades. The Pure Tone Audiometry (PTA) test is one of the most frequently used tests in hearing evaluation and its purpose is to identify if and how well a user can hear sounds at predetermined frequencies. When performing this test, the user is subjected to sounds, each with different frequencies at different volumes, the lowest volume the user can hear a given frequency is considered his threshold and is used to assess his hearing performance. In order to obtain the least tainted results, it is recommended that the user performs this exam in the quietest environment possible, attempting to recreate a soundproof room like environment. [63]

While a normal human ear is able to listen to frequencies between 20 and 20,000 Hz, not all of these are normally tested during a normal PTA test. The most commonly tested frequencies are the following: 250, 500, 1000, 2000, 4000 and 8000 Hz; because this range represents most of the speech spectrum. [82], [83] As such, these were the ones selected to be tested during our game.

To achieve this, first the required technology to reproduce a sound at a given frequency was created and tested. The selected frequency is created dynamically, targeting the left, right or both ears and then is reproduced.

As it was previously mentioned, this prototype should be as self-contained as possible, so it is of utmost importance that a proper earphone/headphone set should be used. The exam would still be performable without it, but doing so will it would incur in a number of problems:

- Forgo the separate evaluation of the left and right ear.
- Required a higher volume output.
- Forgo the noise reduction that is naturally obtained when using an earphone or headphone set..

All of these issues would lead to less accurate results and reduce the effectiveness of the exam.

While earphones sets are cheap and even frequently packaged with a new smartphone, headphones provide a number of advantages over them:

- Higher noise reduction. Even non-noise canceling headphones will do a better job reducing the amount of noise going into a person's ear, by covering the ears in better way than simple earphones.
- Constant placing. When using earphones, it is more difficulty to place them in the same position each time. Sometimes they are placed deeper inside the ear, sometimes the user might feel more comfortable with a slightly looser placement; this will impact the results even if the user always uses the same set. On the other hand, by using a headphone set correctly, the same positions and effect will always be achieved.

For these reasons, when taking one of the auditions exams, it is advised that the user does so while using a headphone set. For a better comparison between exam results, it is also recommended the use of the same set in each test run, thus reducing any possible impact that different sets’ characteristics may have on the results.

With the objective of resembling the traditional PTA test, the game implemented follows its procedure. A sound will be presented and the user should provide feedback that he can hear it. This feedback is in the form of pressing an animal presented to the user, until the sound can no longer be heard. An example of the screen presented to the user can be seen in Figure 23. It is important to notice that the animal and the blue square will always be present, no matter if a sound is present or not. This choice was made in order to not influence the user's response, that is: the user only way of truly knowing if a sound is being presented, is being able to hear it, the image gives no clue to its presence or absence.

Figure 23. PTA Exam (Final Version)

Additionally, to try to prevent conscious or unconscious dishonest answers by the user, no sound (which will now be called zero sound) will be played at random times. This soundless gaps share the same time interval as the actual sounds. Any time the user gives feedback to listening to some noise while no sound is playing, that will be recorded and can be later used to analyze the reliability of that exam.

The way the exam works is: all frequencies start on the lowest volume; upon playing a sound, if the user fails to identify it, that frequency will be added to the pool of remaining frequencies, but at a higher volume. The selection of the next frequency to play is randomized, meaning that a user may hear a sound at another frequency, the same frequency at a higher volume or a zero sound. The user will never receive two zero sounds in a row but may receive two consecutive sounds of the same frequency, but with different volume (assuming he failed to identify the lower volume). Upon identifying a frequency at a given volume, since all frequencies start at the lowest volume, it is recorded as the lowest volume required for a user to hear that frequency, and is removed from the pool.

After the minimum required volume is found for each testing frequency, an audiogram is generated (using the AChartEngine library) and presented to the user, as well as a small summary of his performance and possible problems. The example of an audiogram generated from one trial can be found on Figure 24 and Figure 25. It can be seen the user required a high increase in volume in order to hear sounds at frequencies of 2000 and 4000 Hz and was unable to hear sounds at any volume at 8000 Hz. This would be the final version of the game for this prototype.

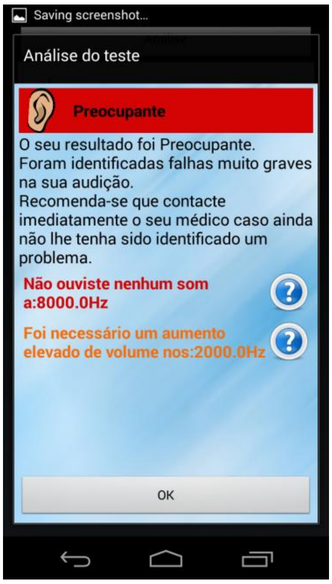

Figure 24. Audiogram Evaluation

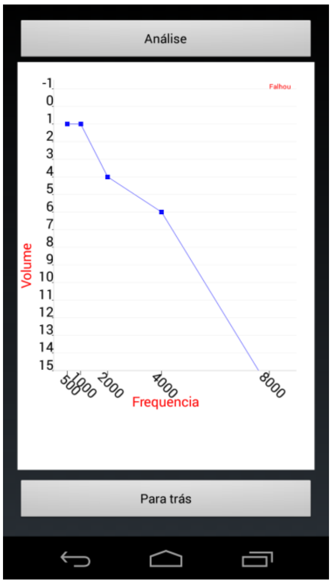

Figure 25. Audiogram

In order to generate and play the sound at the desired frequency, the AudioTrack class from Android's Media API, is used. The exam interface is obtained through the use of standard layout elements (i.e. Image Views). For answer detection, a touch listener is placed upon the square surrounding the animal, which upon being pressed, records the frequency and volume being played.

3.5.6 Speech Discrimination

In the Speech Discrimination test, like reported in Section 2.1.2.2-Speech Discrimination, the user is tasked with listening to a word and then identifying it. The final score of the user is based on the percentage of correct answers. For this exam, various groups of four words were created, each group containing double syllable words [84]–[87] that resemble each other with slight variations. The user will go through the 20 groups, for each group one of the words will be played to the user, through a headphone set like mentioned before, but this time there won't be a left/right ear differentiation. The order in which the groups are tested is randomized. Additionally, for each group, the word to be reproduced is randomly selected by the system, thus making each exam different by both randomizing group ordering and the word being played.

One important matter that was analyzed when developing the exam was whether the selected word should be accompanied by a carrier sentence or not. A carrier sentence is a sentence which envelops the test word, for example: "Say the word ____" or "Repeat the word ____”. In our research, articles were found that both indicated that this helps to obtain more accurate test results [51], [52] and that it has little to no effect on the outcome of the test. [53] Nevertheless, being the worst case scenario a lack of any effect on the test, the choice was made to use the carrier phrase to make the exam more intuitive and comfortable for the user.

Unlike in the PTA exam, the aim of this exam is not to find the minimum volume required. As such, the user should take the exam at his Most Comfortable Level (MCL). In order to find it, a calibration is required. If it is not set before the exam, upon selecting the game the user will be asked to perform the calibration and then will be brought back to the game. The calibration activity is a simple task, based on the methods researched: [54], [88] the user is presented with a sound and he is asked if he would prefer to turn the volume up, down or to remain the same. The user is confronted with the sentence “If this was a television, would you turn the volume louder or lower?” (Via audio) and with the interface instructions below (Figure 26), this explains what the objective is and offers the user with a lengthy sentence to verify his preference.

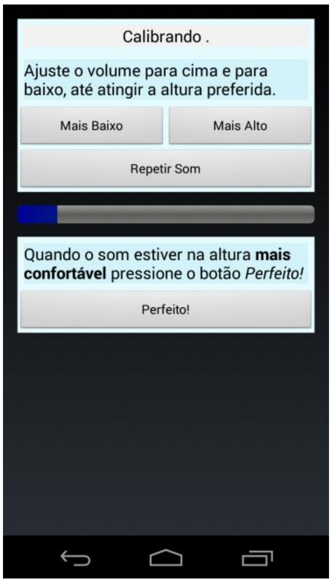

Figure 26. MCL Calibration Screen

In order to create a game with a familiar appearance, the user will play the game with an interface resembling that of the well-known show “Quem quer ser Milionário”, which has been featured in Portuguese prime time television since 2000. As one can see in Figure 27, he is faced with a simple instruction (“Please select the word you heard:”) and given four options.

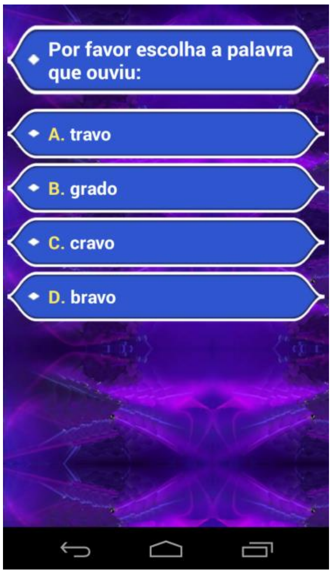

Figure 27. Speech Discrimination Trial

Upon selecting an answer he will go to a money ladder where he will see his current earnings go up or down depending on whether he answered correctly or not. This animation was purposely kept short to not overstay his welcome to the user. Additionally, unlike in the TV show, the user is never disqualified for a wrong answer; he simply goes down one level in the ladder.

The selected words are reproduced to the user through the use of the Android's default Text-to-Speech API. However, due to dictionary availability limitations, the output language is Portuguese with a Brazilian accent. The ladder animations are a combination of Image Views and Runnables, which alter arrow position in relation to the other elements of the layout.

Upon completing the trials the user is present with a percentage based score and with a small evaluation of his result. The grading table [89], [90] used for the qualitative evaluation given can be seen below (Table 2):

Table 2. Speech Discrimination Evaluation

Score Range	Evaluation
89 ≤ Score ≤ 100	Excellent
77 ≤ Score ≤ 88	Good or slight difficulty
65 ≤ Score ≤ 76	Fair to moderate difficulty
54 ≤ Score ≤ 64	Poor or great difficulty
Score ≤53	Very poor

In order to obtain the least tainted results it is recommended that the user performs this exam in the quietest environment possible (trying to achieve a soundproof room like environment) and with the help of the same headphone set between test runs, so that if a result comparison is made, it can be more accurate.

3.6 Other Features

Some features that are not games but that help in the testing process will be described here.

3.6.1 Light Sensor

The light sensor normally available in smartphone devices will be used to make an assessment of the environment the user is about to perform the exam in. For this, Android's sensor API will be used. Upon choosing to perform an exam, this sensor will be used to evaluate the amount of light the device is faced against. At high levels it might indicate the user is taking the exam at unfavorable conditions, thus possibly having a negative effect in test performance. If the obtained value (in lux) is higher than 1000 (the values associated with overcast day light and bright indoor environments [91], [92]), then the user is advised against taking exam and warned that if he continues, the results might be lower than expected.

3.6.2 Distance Measurement

To aid the user in placing the device at the desired distance for vision exams (in order to keep the exam result accurate and comparable between test runs), a distance measurement tool was developed. This tool uses OpenCV to analyze images captured through the device's frontal camera, after analyzing them it is able to retrieve the user's eye distance. By using triangle similarity (as shown in Figure 28), Formula 7 is achieved, which can be converted to Formula 8 and, with the information collected (pre-determined device-user distance, image eye distance and user inputted real eye distance), it is possible to perform a calibration of the focal length of this device.

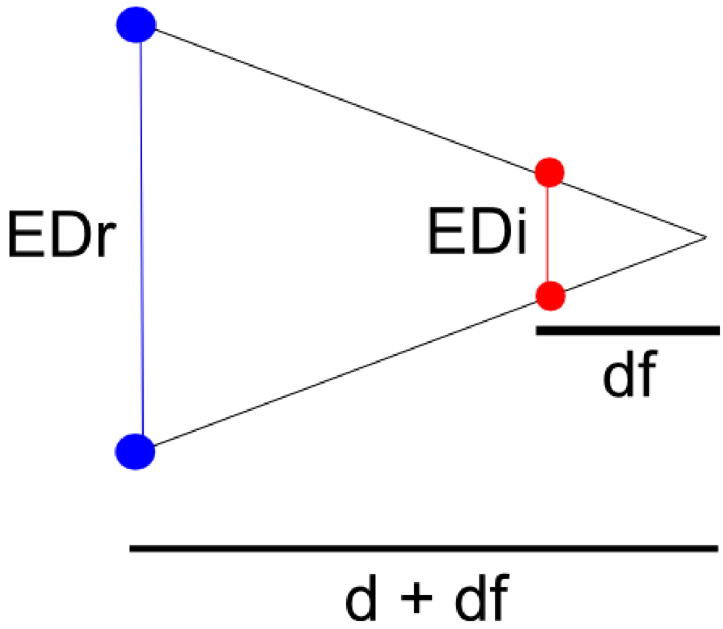

Figure 28. Eye distance measurement

$$\frac{EDi}{df} = \frac{EDr}{d+df} \qquad (7)$$

$$df = \frac{EDi * d}{EDr - EDi} \qquad (8)$$

In the formulas above, *EDi* and *EDr* represents Eye Distance in the **i**mage and in **r**eality, respectively, *d* represents device-user distance and *df* is the focal length.

Later, when the user performs his vision exams, by referring to the Formula 9 (also obtained from Formula 7), he can use this tool to calculate the user to device distance. With this information the prototype can give feedback to the user to adjust his position accordingly.

$$d = \frac{(EDr - EDi) * df}{EDi} \qquad (9)$$

The testing conducted in the development phase of the system seemed to indicate an ability to get an approximate value of the user's distance, which might prove useful to the test procedure. However, due to the priority being in the development of exams, the reliability of this tool's results was not able to be fully tested. Nevertheless, after conducting some testing with volunteers, it was possible to assess this tool's importance, which will be discussed in the conclusion section of Chapter 4 - User testing and Validation (53). It must be noted that this calibration process is possible but relatively difficult to perform without help, however, this is a onetime calibration, thus help, if required, is only needed once.

3.7 Main Interfaces

Upon starting the application, the user is presented to what will be called the main screen. This screen presents the user with two simple options: either enter the “Jogos” (“Games”) or “Outros” (“Others”). Most of the user’s actions will hopefully be the playing of games, and the occasional visit to the Results or Options menu, thus every non-game activity was confined to the “Others” section.

Figure 29. Main Screen

Continuing to provide the user with a small number of choices at a time, after selecting “Games” the user is taken to a screen where he must choose between Vision (“Visão”) or Audition (“Audição”) games, or to Go Back (“Para trás”) to the previous menu.

Figure 30. Games Screen

Upon selecting the chosen game type, the user is taken to a screen with the games options available for that sense. If he opts for the Audition games, he will be presented with a similar layout containing two games and a message telling him to choose the game he wants to play.

Figure 31. Audition Games Screen

If he opts for the Vision games, a slightly different layout is provided. Here he is presented with a message to choose which games he wants, and with a grid of games amongst which to choose from.

Figure 32. Vision Games Screen

Upon completing the game, he is presented with a quantitative and a qualitative evaluation of his result. The user can use this feedback to monitor his performance, taking action if he suspects that there is a problem with his senses. The data for each exam is always stored for easy comparison in the Results menu, which can be accessed by selecting the "Others" options (seen on Figure 33) instead of "Games" in the Main Screen (as seen previously on Figure 29), as seen in the image below.

Figure 33. "Others" Screen

Throughout all this process the user is always presented with a small number of clearly labeled buttons and nothing more is expected of him than a click. If he ever wants to go back to where he came from, he simply clicks the "Go back" button available. This button was created to avoid the need to use the Android's default back button, thus keeping all required interactions from the user within the application interface. Upon returning to the main screen, another go back instruction would make him leave the application, to signal this, the button is changed to a red color and the text changed from "Go Back" to "Exit". If the user clicks the "Exit" button a dialog will be shown asking him to confirm his decision.

Figure 34. Exit Confirmation Screen

3.8 Summary

This chapter focused on detailing the development process. Initially the test selection process was covered, explaining why each test was chosen and why others had to be discarded. Following that, considerations that were taken into account when envisioning the test and game procedures, in order to produce simple, quick but effective games, were exposed. Finally, a detailed view of each exam development process was given, explaining exam-specific considerations, presenting different iterations of the development and different ideas that were considered when deciding test and game procedure. Additionally, general and game specific interface considerations and appearance were mentioned, as well as small explanations of how they were made.

After this system was developed a validation of the final prototype was needed. To do this a sample of volunteers from our target users was required. The next chapter will detail the test procedure that was followed, as well as conclusions drawn from the process and an evaluation of the results.

Chapter 4

User testing and Validation

This chapter details the testing procedure in a number of volunteers. It also presents results, considerations and discussion about details gathered during the tests and upon result analysis.

4.1 Overview

Upon contemplating which exams should be performed by the user during the testing sessions, a decision was made to perform all of them. This would allow us to validate the results in some of exams and the usability in others. With the information collected and observed during the tests an analysis was made to evaluate possible improvements to the work.

Thanks to partnership already established by Fraunhofer Portugal, a number of participants that were willing to be testers of the system were contacted and recruited. Prior to validation, each user was briefly informed about the aim of the tests and inquired about possible vision or audition deficiencies. No sort of compensation was offered to the volunteers in this study.

Additionally, in an attempt to validate some of the exams, some reading charts that the user would be submitted to were acquired, so that a correlation between the results obtained using the prototype and the ones from the charts could be analyzed.

The tests were conducted in Fraunhofer Portugal's Living Lab, in an area specially designed for prototype testing which resembles a small apartment. This provided a good environment in which to test the prototype, especially considering that this is the intended environment for the user to conduct his exams in. Still, it presented some limitations, mainly that mirrors were only present in a small bathroom; however, this space was still enough to conduct the Distance Visual Acuity exam.

4.2 The Volunteers

Ten different individuals were tested. Prior to starting the testing process the user would be inquired about known vision or audition deficiencies. If one was acknowledged, the user would also be asked for further information, such as: severity and which side was worse. If the volunteer was able to answer any of these questions, this information would be recorded.

The volunteer's age ranged between 64 and 76, with an average age of 70.5 and standard deviation of 3.7. This sample allowed us to perform exams on a group of people that was similar to the target audience of any future release of a prototype based on this work.

The users' experience with mobile devices was all rather small. Actually, all but one of the users stated that they did not own a smartphone or similar touch based device, or even used such device on a frequent basis. However, all of the volunteers used a smartphone device at least once, most of them in similar test-based environment.

More detailed information about the users which performed these tests can be found in Appendix A-Volunteers Information.

4.3 The Paper Tests

While it was not possible to reproduce the traditional audition tests used in a clinical setting, due to the lack of the equipment required in order to be performed correctly, some of the vision tests were easier to reproduce. As such, in order to validate the results obtained from the prototype, an attempt to reproduce the traditional paper-based vision tests was made. These charts and cards were obtained after contacting manufacturers, distributors and clinics. From these contacts, the items that were possible to acquire were the following three vision charts, which were then used to test each volunteer:

• Visual Acuity Chart – the Oculus' "Illiterate E's Visual Acuity Chart For Distance", with reference number 4692, was used. This chart measures distance visual acuity and was intended to be used at 5 or 6 meters. The chart is built in a similar way to the Snellen chart but, instead of using letters as its optotype, it features the Tumbling E. The user was tasked with correctly identifying each optotype's orientation. The chart was placed on a wall and the user stood at a distance of 5 meters from it. The chart was composed of ten lines, each with a visual acuity reference value, being the topmost line the equivalent of a visual acuity score of 0.1 and the last one with 1.0 (the higher the score, the better), the increase in score was 0.1 per line. When the user was no longer able to answer correctly to the majority of the optotypes in a line, the test was stopped and his score was recorded, based on the last read line. Even though most volunteers would be reluctant to answer when they were not sure of the answer, they were encouraged to make a guess based on what they thought was correct. This revealed that most users could actually achieve a better result than they seemed able to at first. Figure 35 shows the chart used.

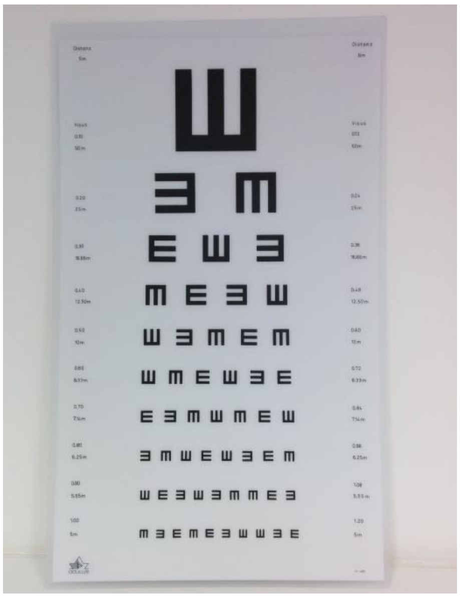

Figure 35. Oculus' Illiterate E's Visual Acuity Chart For Distance

• (Near Vision) Reading Chart - A ZEISS Portuguese reading chart was used. This chart is intended for use at a 40 centimeters distance. This chart was used to reproduce a reading test and featured eight possible scores for visual acuity: 0.2, 0.32, 0.4, 0.5, 0.63, 0.8 and 1.0 (the higher the score, the better). On the smallest letter sizes, multiple paragraphs were present, helping to reduce memorization bias between test repetitions. Depending on the smallest line the user could read, he would be given the score for that paragraph. To achieve the desired distance a measuring tape would be used to place the chart 40 centimeters from the user's eyes and would be held there by the tester. The user would be instructed not to alter the distance to the chart, leading to a null test if he did. Figure 36 shows the chart used.

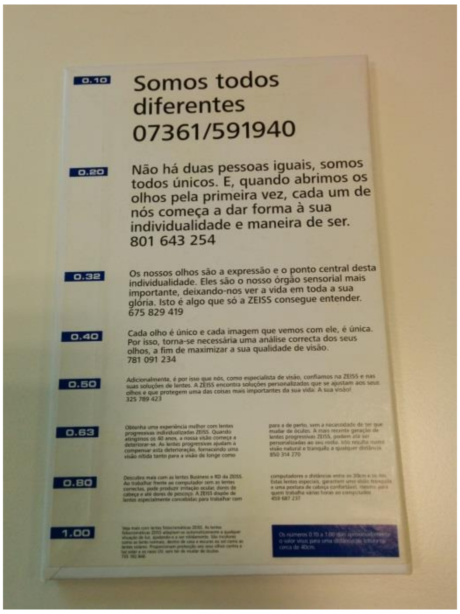

Figure 36. ZEISS' Portuguese Reading card

• (Near Vision) Visual Acuity Chart – a Runge Pocket Near Vision card, distributed by Bausch + Lomb and Precision Vision, was used. This chart is intended for use at 40 centimeters and an identical process to the reading chart was used to assure the correct test distance. It features sixteen columns with three rows per columns. The optotypes used in this card are the Sloan Letters. The volunteer was tasked with reading each column and then proceed to the next. Upon being unable to correctly identify the letters in a column, his score would be set. A picture of the card can be seen on Figure 37. In the front (the side visible) is the test which the users sees for testing purposes. In the back (not visible in the picture) there is the reference, where the tester can verify the answers given and the final score to assign to the user.

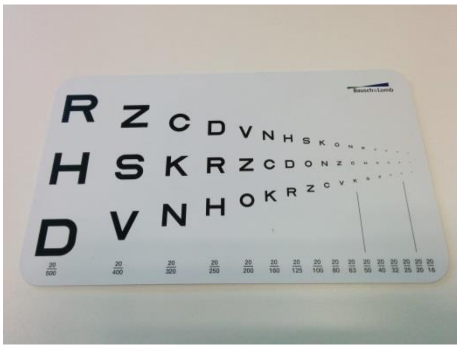

Figure 37. Runge Pocket Near Vision card

4.4 Test Procedure

In total, a complete test run was comprised of 9 tests: 6 prototype based (exams) and 3 chart based. If possible the user would take every test, with both sides (eye/ear, when applicable) and with or without aid (glasses/hearing aid, when applicable).

The testing procedure would be carried on in the following order:

1. Briefing of the dissertation's goal;
2. User experience and vision and audition quality inquiries;
3. Chart based testing;
4. Prototype based testing;
5. Usability questionnaire and suggestions phase.

During Phase 1 the user would be re-briefed on what the dissertation's objective was and given a small summary of the process he was about to start. After this explanation, the volunteers were presented with a form giving their consent to record this test procedure, with the knowledge that only the person conducting the test would have access to it and only in the context of this dissertation.

Phase 2 was used to gather further volunteer data. The user's experience with smartphones and context of those experiences were analyzed. They were also questioned about possible vision or audition problems, such as: present known problems, use of (vision or hearing) aid and past problems that might have been corrected.

Phase 3 was comprised of chart based testing. Each user would perform the three tests listed in 4.3 -The Paper Tests. If in step 2 the user identified a worse eye, that would be the starting eye for each test.

Both phases 3 and 4 required the user to perform vision tests, with one eye at a time. To achieve this, the user would use an eye patch that would be placed on the other eye (the one which was not being tested). When required, the user would be provided with help so that the eye patch placement did not influence the tested eye due to incorrect placement. In phase 3, if one eye was recognized as worse than the other, he would always use the worst eye first, in an attempt to reduce test repetition bias. If the tests were being conducted with and without glasses, they would first be performed without them (with both eyes separately) and then with them (again, with both eyes separately).

Phase 4 was when the prototype was tested by the user. The user would be tasked with performing a set of tasks and their performance would be recorded. Initially the tasks that the user was to perform were:

1. Go to “Jogos de Visão” Menu

Ideal Flow: [Menu Principal -> “Jogos” -> “Visão”]

2. Select and Perform “Visão ao Perto” exam [Prior to selecting the exam, a brief explanation was given].

Ideal Flow [Already in “Visão” -> “Visão ao Perto”]

3. Select and Perform “Contraste” exam [Prior to selecting the exam, a brief explanation was given].

Ideal Flow [Already in “Visão” -> Contraste]

4. Go to “Jogos de Audição” Menu

Ideal Flow: [Already in “Visão”-> “Para trás” -> “Audição”]

5. Select and Perform “Identificação de Palavras” exam [Prior to selecting the exam, a brief explanation was given].

Ideal Flow [Already in “Audição” -> “Identificação de Palavras”]

6. Select and Perform “Deteção de Sons” exam [Prior to selecting the exam, a brief explanation was given].

Ideal Flow [Already in “Audição” –> “Deteção de Sons”]

7. Go to “Resultados” Menu

Ideal Flow: [Already in “Audição”-> “Para trás” -> “Para trás” → “Outros” → “Resultados”]

8. Go to “Jogos de Visão” Menu

Ideal Flow: [Already in “Resultados”-> “Para trás” > “Para trás” -> “Jogos” → “Visão”]

9. Select and Perform “Comparação de Formas” exam [Prior to selecting the exam, a brief explanation was given].

Ideal Flow [Already in “Visão” -> “Comparação de Formas”]

10. Select and Perform “Visão ao Longe” exam [Prior to selecting the exam, a brief explanation was given].

Ideal Flow [Already in “Visão” -> “Visão ao Longe”]

Before starting to perform the task process above, the user was given a time window to freely use the prototype, in order to get accustomed to the touching interface and the general interface flow. Upon asking the user to perform a game based task (exam), the user would be given a small explanation as to what its objective was and what was expected of him (game interaction wise). In order to give each user the same degree of information about the exam, a standardize amount of guidelines were set and were then read to the user, these can be found in Appendix A-Task Specific Guidelines.

Because the user performs Task 5 using both ears at the same time, it was the only exam that was performed only once (baring normal test conditions). The remaining game based tasks (2, 3, 6, 9 and 10), were all required to be taken more than once. This is due to the fact that, when possible, the user should take the exam using the left and right eye or ear, and with and without vision or hearing aid.

On phase 5, the user was asked for opinions regarding the prototype and the exams he had just taken and given a System Usability Scale questionnaire to further record his opinion.

4.5 Exam Tasks Procedure

This section details the test procedure followed during the exam related tasks. It explains the steps taken in order to keep results accurate and maintain the same testing conditions between volunteers.

4.5.1 Vision Games

All vision exam, except the Distance Visual Acuity one (Task 10), followed the same procedure:

- The user is sitting down at a table
- The user is informed the task he is about to perform
- The user receives a small explanation of what he is to do once the task starts
- The user is informed to perform said task, using a specific eye.

The user was asked to perform the exam, first without glasses, using the worst eye first if identified, otherwise starting with the right eye. The non-tested eye was covered with the previously mentioned eye patch (as an example, Figure 38 show the one used in these tests). If applicable, the user would then repeat the exam with each eye, but this time using his glasses. Not all exams needed to be performed four times, because most users had glasses that were only for near vision or distance vision, not both.

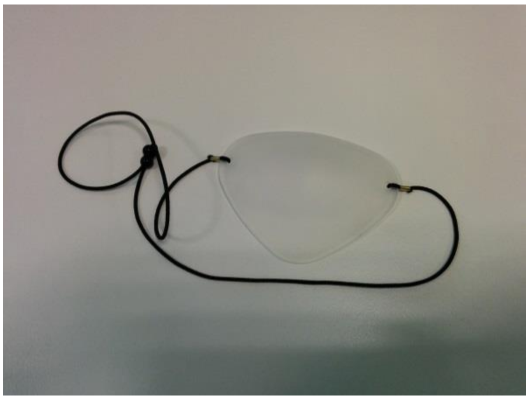

Figure 38. Eye patch

Upon selecting the exam to perform and the eye to use, the timer was stopped and the user-device distance was set to 40 centimeters. To do this, the tester used a measuring tape to set the distance between the eyes of the user and the device (at a comfortable position, which he was to maintain during the exam). In order to reduce distance variation, an object was placed on top of the table. The user should have the device held against the object and he should try not to reduce the distance between himself and the device. The user would then proceed with the exam and the timer would be resumed. During the game progress, it was frequent to make re-measures of the distance in order to alert if it had decreased. This repetition measurement was conducted in such a way as not to interfere with the user.

The Distance Visual Acuity (Task 10) is the only exam which was required to have the user in a standing position. This was due to the environment limitations which wouldn't allow the exam to be taken in a sitting down position. Approaching the reality of an actual house, the mirror on the Living Lab was used and, with limited space, it would be impossible to use a chair without compromising the test distance. This limitation might or might not be present on a potential user's household. For this exam, a mark was drawn on the floor at a distance of 1.5 meters from the mirror, thus mimicking a 3 meters visual acuity test. The user was asked to start the exam and turn the screen to face the mirror and hold it against his body. This way, the user would only be able to see what was on the device's screen by looking at the reflex on the mirror. The user's objective was to physically rotate the device in his hands, until the optotype (a Tumbling E, the same used in the chart based testing) had its leg pointing to the ceiling, at which point he should make a long press on the screen.

4.5.2 Audition Games

There were two distinct tasks related to Audition games: the Speech Discrimination game and the "Sound Detection" (Pure Tone Audiometry) game. The former required a prior calibration of the user's Most Comfortable hearing Level (MCL) to be performed. This calibration was performed before the task step was initiated. Both of these exams required the user to wear a headphone set. The set chosen was a Sennheiser HD 201 (seen on Figure 39), for its ability to reproduce the desired frequencies and because it is a set that would be easily available for the users to acquire, thus being similar to a future situation where the user takes advantage of the system on his own.

Figure 39. Sennheiser HD 201 headphone set

Like reported previously, the Speech Discrimination exam reproduces one word (inside a carrier sentence) to the user at his MCL and he is then tasked with identifying it between four possible options. This exam is taken with both ears at the same time.

The Sound Detection, similar to the vision exams, required the user to perform it one ear at a time. However, unlike the vision exams, here the user simply selects which ear he is testing and the sound will only be reproduced on that side. In the case of hearing aid usage, the user was asked to take the exams with and without the aid.

4.6 Evaluation

By comparing the data gathered from the users during the introductory questionnaire and the chart based tests with the prototype tests, an analysis of the results obtained was made. From this comparison it was possible to evaluate a correlation between the tests results and to find ways in which the prototype developed could be improved.

Before continuing there is one topic that needs to be discussed, and that is: Testing time. Unlike the intended use of the prototype, where the user performs these games when he wants and in any sequence, this testing procedure forced the user to perform a large number of tasks in a row (being prototype based or not). This led to a big testing time which was noticed to have an impact on the overall performance of these elderly users. All the volunteers are of advanced age and bring a large set of adversities that are not found on the average smartphone user, being: low experience with the devices, limited motor skills (hand or eye shaking, difficulty standing up, difficulty rotating the device) and a strong impact of continuously attempting to do their best on the tests. This would reflect in different degrees on most people, but the ones that were on the most troublesome end, either required small intervals between tests or additional help when performing tasks. Due to this, after a considerable number of tests were performed and features that could be improved were identified, some tasks that were not related to sense evaluation started to be discarded, in order to reduce discomfort related to testing time on the volunteer.

4.6.1 Interface Evaluation

This section focuses on detailing the interface evaluation obtained from the users' interaction with the device.

The interface related tasks that were asked of the user were the following:

- (Task 1) Go to "Jogos de Visão" Menu (From the Main Page).
- (Task 4) Go to "Jogos de Audição" Menu (From "Jogos de Visão).
- (Task 7) Go to "Resultados" Menu (From the "Jogos de Visão").
- (Task 8) Go to "Jogos de Visão" Menu (From "Resultados")

Different users performed these tasks with different time and assistance required. Task 1 and 4, which asked the user to go to different sense related Games, required the expected amount of time, leading us to believe that they were easy to reach, which is important since this is the main focus of the system usage.

Task 7 would take a little more, yet still expected amount of time. This is probably explained by the fact that it was the first time most users would go into the "Others" options.

From observation of the users' interaction one could see that the most complex of these tasks was Task 8 (returning to Vision Games after going to Results). This might be due to a combination of two options:

- The way the task was formulated. Upon asking the user to go to the "Results" menu and then immediately asking him to go to "Vision Games", this might have generated some confusion in understanding what was really intended of these tasks. The user had already been (and performed) in "Vision Games" and had been sent to the "Results"; once there, no task was asked from him, but was asked to go again to “Vision Games”. The user might have felt that going back would make little sense, especially if there were Vision related options (the exam results) on the Results menu he had just been sent to. This would make it seem like the previous task was pointless (which it in fact was, besides interface testing purpose), perhaps making him reluctant to immediately go back.

- The Results menu interface. Allied with the previous reason, this might explain the sense of confusion in the user. The “Results” menu (Figure 40) is labeled as such and the user was already familiar with the “Vision Games” menu (Figure 30); however, in an attempt to make it more intuitive to the user it is possible that more confusion has been generated. The clickable buttons on the “Results” menu share the same image and name as their “Vision/Audition Games” counterpart, in an attempt to ease the user’s task (if he wanted to check exam results, it would be easy for him to identify which option to choose). However, when asked to go to the “Vision Games” Menu, seeing options such as “Visão ao Perto” (“Near Vision”) and “Visão ao Longe” (“Distance Vision”), the user would normally click one of those options, in an attempt that this was the objective, which was not.

Figure 40. Results Screen

This leads us to believe that, in order to provide the user with an easier workflow, design changes should be adopted. Potential, changes are: make the menu interface different from the previous menus, icon differentiation and clearer renaming of the options. All of this in an attempt of expressing that it is actually the results and not the games that are being presented. It is also believed that the way the instructions were given to the user, generated confusion that wouldn’t be present if the user had made the option of pursuing the Results’ himself (in an attempt of checking a exam result).

4.6.2 Results Comparison

In this section, the exam results obtained using the system are compared against the paper based tests and the data gathered from each volunteer, attempting to find correlations between them.

4.6.2.1 Distance Visual Acuity

In this, the Distance Visual Acuity test will be analyzed. In Figure 41 it is represented the graph generated from 22 test results (some outliers were discarded); the blue line represents the results obtained from the chart based testing and the red line represents the results obtained from the prototype tests.

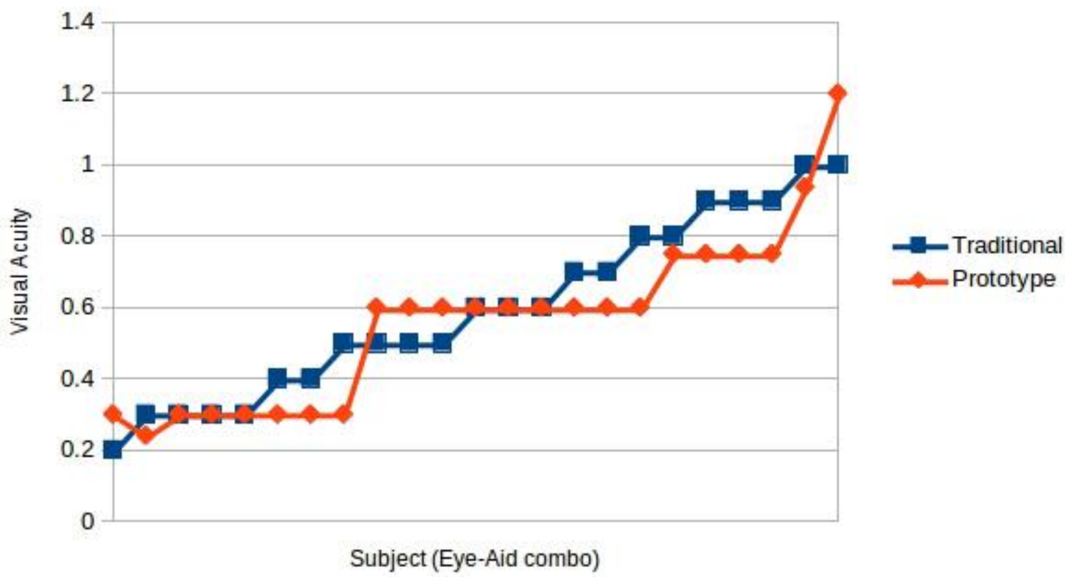

Figure 41. Distance Visual Acuity Results

The vertical axis represents visual acuity and the horizontal axis represents a test eye-aid combination of one user. The points in the blue line, represent the score obtained using the traditional charts and in the red line, represent the score obtained in the prototype, by the same eye of the user. Each point represents one test trial performed on one eye, and the set of points was ordered in ascending order of traditional score, to produce the line graph seen.

While not being an exact match, the results obtained from the traditional method and from the prototype seem to have a good correlation between them. There are a number of factors that can impact the result of either test, two of them that were are:

- Fatigue – in some users, eye shaking and general fatigue was noticed when performing the large number of tasks (like mentioned previously) that were required in this validation procedure. The focusing on one sense, with repetitive tasks that encouraged the user to force their vision, seemed to produce a sort of exhaustion of the eyes.

- Dealing with uncertainty – this is another factor that was noticeable when conducting the initial chart based testing. Some users were reluctant to give their answer on some optotypes if they weren't 100% certain it was the correct one (even though when encouraged to do so they would, for the most part, obtain a better result). Some people were more reluctant than others and in the prototype based testing far less assistance was given, thus perhaps, reducing the user’s commitment to make an educated guess.

These might account for some of the changes that are perceived in the graph above. Nevertheless, the exam seems to be able to obtain a good approximation of the user's visual acuity which is useful in detecting alarmingly low values and to track sudden drops in performance. A similar test is hard to find, and by enabling the user to perform it on their own and with little investment, this solution may be a powerful tool in vision assessment between clinic visits.

4.6.2.2 Near Vision

The following table lists the results obtained during the reading and near visual acuity tests performed using the cards and the system reading exam.

Table 3. Near Vision Acuity Tests Results

User Code	Eye	Aid	Reading VA	Letter VA	Prototype
0002	Right	Yes	0,4	0,93	0,4
0002	Left	Yes	1	0,8	0,5
0003	Right	Yes	ND	0,625	0,5
0003	Right	No	ND	0,25	0,25
0003	Left	Yes	ND	0,2	0,16
0003	Left	No	ND	0,125	0,1
0004	Right	Yes	1	0,4	0,4
0004	Right	No	0,5	0,3	0,25
0004	Left	Yes	0,63	0,3	0,32
0004	Left	No	0,4	0,3	0,125
0005	Right	Yes	ND	0,3	0,4
0005	Right	No	ND	0,3	0,25
0005	Left	Yes	ND	0,3	0,32
0005	Left	No	ND	0,3	0,125
0006	Right	Yes	1	0,5	0,25
0006	Right	No	0,4	0,5	0,25
0006	Left	Yes	1	1	0,5
0006	Left	No	0,4	0,5	0,25
0007	Right	Yes	0,5	0,4	0,25
0007	Right	No	0,63	0,3	0,32
0007	Left	Yes	0,5	0,5	0,32
0007	Left	No	0,63	0,4	0,25
0008	Right	No	0,4	0,5	0,25
0008	Left	No	0,4	0,5	0,32
0009	Right	Yes	0,8	0,8	0,5
0009	Right	No	0,4	0,3	0,2
0009	Left	Yes	0,32	0,3	0,125
0009	Left	No	0,2	0,2	0,1
0010	Right	Yes	1	0,8	0,5
0010	Left	Yes	1	0,8	0,5
0011	Left	Yes	0,8	0,625	0,32
0011	Left	No	0,63	0,625	0,32

The table above represents the Near Vision results obtained on both the traditional tests and the prototype test. The first column indicates which user this row belongs to. The second column indicates whether the eye used was the Left or the Right one and the third one whether the user performed the test with vision aids or not ("Yes" or "No"). The fourth, fifth and sixth columns indicate the visual acuity score obtained in the traditional reading test, traditional letter test and the in prototype reading test, respectively. It is important to notice that, even between both traditional tests (Reading and Letter acuity), there are some very different values for the same user.

Unfortunately, this provided some highly fluctuating values as can be seen in Figure 42 and Figure 43, comparing Reading and Letter test acuity against the prototype results, respectively. When considering these graphs, it is important to notice that because of the device-distance combo limitation, the traditional scores that exceed the value of 0.5 (the upper limit for this exam in the device used) were replaced with 0.5. Actually, prototype tests that reached the 0.5 would be stopped and marked as inconclusive, because it was a limitation of the device, and not the user. So, these high values (over 0.5) were replaced with the value 0.5 when creating the comparison graphs (Figure 42, Figure 43 and Figure 44), for an easier comparison of results.

Below, the chart obtained from comparing the scores obtained in the traditional reading test (blue line) and in the prototype test (red line) is visible. The vertical axis represents visual acuity score, while the horizontal axis indicates an eye-aid combination.

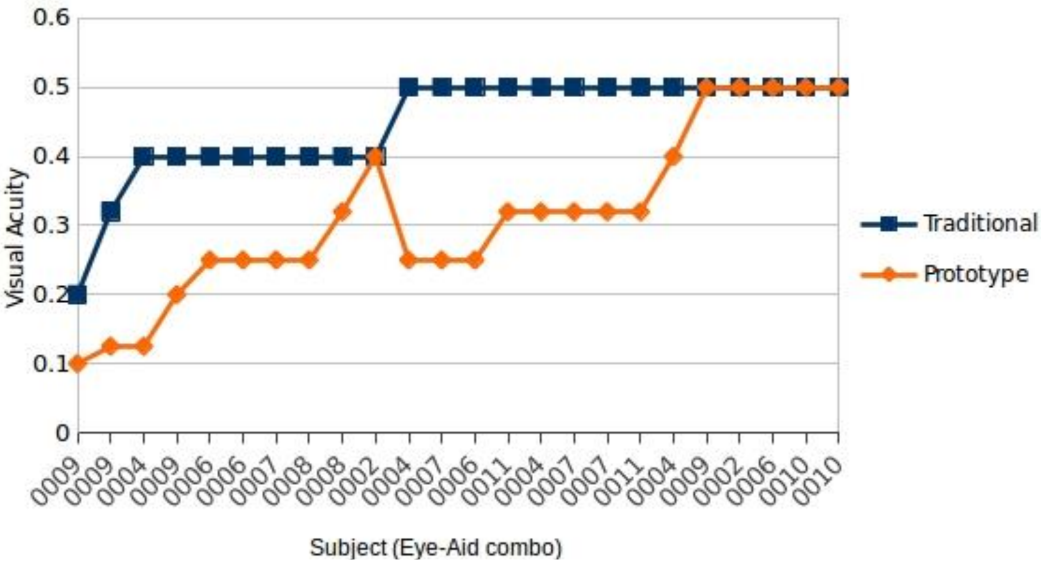

Figure 42. Reading - Prototype Visual Acuity Comparison

From analyzing Figure 42 it is possible to see that the results obtained from the exam are constantly below the traditional method score. The only exception is for some 0.5 scores, but those remain inconclusive because the user reached the device's limit and could be equally different. However, there is another factor worth considering: the test procedure is not an exact match and this most likely has an effect on results obtained. The biggest change is that, whereas in the paper based test the user only has to read the card, in the prototype the user is faced with an extra challenge: finding the missing letters. From the observations of the test subjects, this undoubtedly makes a difference. During testing it was very frequent, to see the user being able to read the sentence almost entirely (meaning his vision was good enough), but to be stuck on a particular word. Sometimes they were unsure about what the word was even when they could read the present letters of the incomplete word; other times they had doubts about the letter that was missing, due to the fact that different letters ended up having different difficulties. An example of this is the letter "l" (small "L"): because this letter is rather thin, if it was placed before or after the occluding square, in smaller letter sizes, it would be very hard to read, meaning the user is now facing a word with two missing letters instead of one. While this problem could be mitigated through some larger spacing or by capitalizing of letters, the major problem would still remain: they are not the same test. So a different exam procedure must be designed, or a similar test in paper form needs to be created and validated, so that the prototype evaluation is not tainted by test differentiation.

The same near vision prototype scores will now be compared against the tradition letter-based visual acuity test. Following the same procedure, Figure 43 was created.

By analyzing Table 3, it can be seen that in some users, great differences occur between the results obtained during traditional letter-based visual acuity testing and prototype testing.

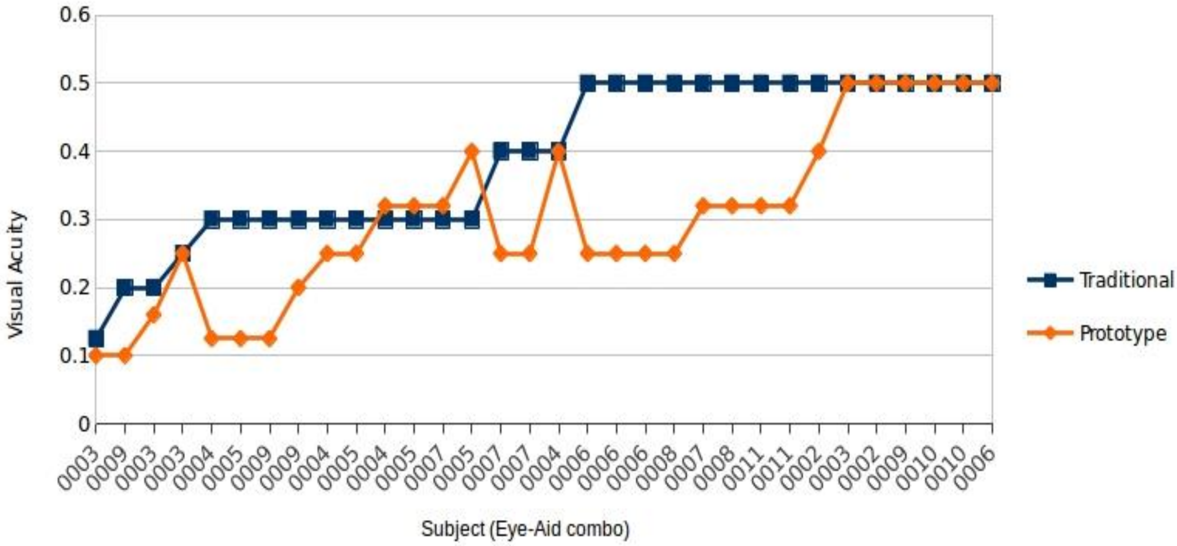

Figure 43. Letter - Prototype Visual Acuity Comparison (Unfiltered)

After removing some of the most critical offenders (0005, 0006, 0007, 0008 and 0011) a better graph is obtained, as seen in Figure 44 (the values have been reordered).

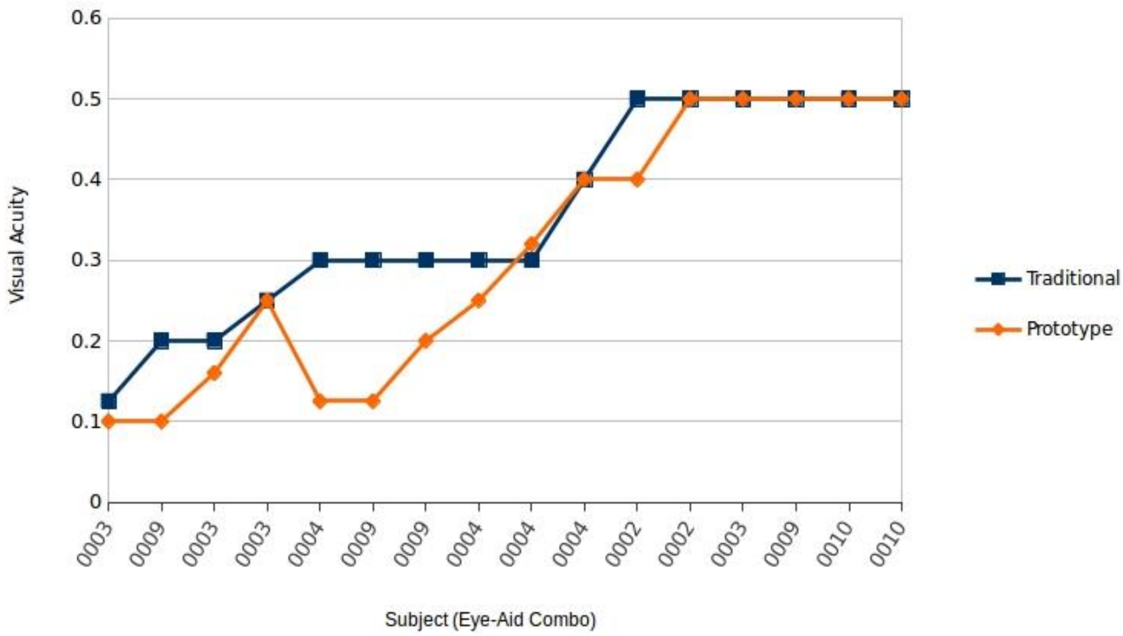

Figure 44. Letter - Prototype Visual Acuity Comparison (Filtered)

This graph still has some outliers, mainly represented by user 0004 and 0009, which scored lower on the prototype test than in the paper test (possibly due to the reasons already mentioned). Nevertheless, a reasonable correlation seems apparent, especially considering that the difference in test procedure may account for some variation.

However, half the test participants had to be removed, meaning that with the sample of users that performed these tests, an accurate evaluation of this test cannot be made. Conclusions about this exam's accuracy and reliability cannot be made and, as such, a larger sample would need to be tested, in a best case scenario, even repeating test subjects and conditions (same eye, same presence or absence of glasses, etc.).

Comparing against the reading test (Figure 42), the exam seems to be reasonably able to distinguish between eyes with different acuities. However, this exam yields different scores than those obtained with the traditional test, because the exam's score is always inferior to the traditional one (as seen previously). Therefore, this exam, in its current state, is not able to pinpoint accurately the reading acuity of a user. It might still be able to allow the user to keep track of any decrease in performance. However, the more accurate the score is, the better an assessment can be made when combining exam results. So, like mentioned before, this exam requires either a different approach or a different comparison test to be validated against.

Comparing against the letter test (Figure 44), while the results showed an ability to make approximate estimations of the user's vision, it did so with too many outliers to be reliable. If further tests were to be carried out (both repeating users and conditions, or with new users) and the percentage of outliers was reduced, this exam could prove itself useful for tracking the development of the user's vision.

So, in summary, the testing done seems to indicate that, while the exam has some potential advantages, in its current form it would still have the short coming of not being able to accurately pinpoint the user's near visual acuity. As such, further work needs to be developed for it to reach a high degree of accuracy.

4.6.2.3 Contrast Sensitivity and Shape Discrimination

Both Contrast Sensitivity and Shape Discrimination exams were performed by the users, using the same procedures of the near vision test: one eye at a time, with and without glasses. These results can be found in Appendix A- 7.2.Volunteer Test Results. However, without the traditional tools to compare these tests against, conclusions cannot be drawn. As such, clinical trials using professional tools and methodologies would have to be conducted so that a comparison of results could be made.

Additionally, based on the sample at hand, significant correlations between the results of these exams and the visual acuity exams could not be made. As such, the impact of these exams on the vision assessment by this system requires further investigation in order to be validated.

4.6.2.4 Audition

Unfortunately, the audition tests are more complex and it was not possible to reproduce them. So, it was not possible to obtain results using the traditional methods, which would allow for an assessment of our exams' accuracy. However, with the information gathered from the users' about their vision and audition it was possible to analyze the exam results and look for a correlation between the users' information and the results obtained in the exams.

Among the ten users tested, four of them expressed known hearing problems and one (with code 0006) expressed the feeling of having a hard time understanding people, even though he was not diagnosed with any problem yet. Among those who reported hearing problems, two were recommended the use of hearing aids but only one of them did. One Pure Tone Audiometry (PTA) per ear and one Speech Discrimination exam per person were performed. Additionally, the volunteer with hearing aid repeated PTA exam, without the aid.

In this section, the results obtained with the prototype (using the procedure explained in 4.5.2) are compared with the information gathered from the users. The results from the users which declared good hearing are examined for similarities. The results from volunteers which declared faulty hearing are compared against normal hearing ones, and are examined separately.

The following is a table of the volunteers' results.

Table 4. Complete Audition Tests Results

Code	250	500	1000	2000	4000	8000	Speech Disc. Score	Problem
0002 - right	2	1	1	1	3	4	95%	No
0002 - left	1	1	1	1	1	6		
0003 - right	3	1	1	1	2	6	95%	Yes
0003 - left	1	1	1	1	3	8		
0004 - right	2	1	2	1	1	2	100%	No
0004 - left	2	1	1	1	1	1		
0005 - right	2	1	1	1	1	1	95%	No
0005 - left	1	1	1	1	1	1		
0006 - right	2	1	2	1	5	3	85%	Maybe
0006 - left	1	1	1	2	3	3		
0007 – right (with aid)	1	1	1	3	2	6	95%	Yes
0007 - left	7	7	6	8	X	7		
0007 – right (without aid)	1	1	1	7	8	8		
0008 - right	5	4	3	2	7	X	80%	Yes
0008 - left	1	1	1	5	5	X		
0009 - right	1	1	1	1	1	1	95%	No
0009 - left	1	1	1	1	1	2		
0010 - right	1	1	2	1	1	3	100%	No
0010 - left	1	1	1	1	1	4		
0011 - right	2	1	1	3	6	X	85 %	Yes
0011 - left	2	1	1	3	4	X		

Table 4 represents the results obtained from 21 PTA and 10 Speech Discrimination exams that were carried out. The columns are as follow:

- Column 1 is the user's identification code. It also identifies the ear used for this exam. For user 0007 it also mentions whether he used hearing aid or not; all other subjects performed the exam without aid.
- Columns 2 through 7 (250-8000 Hz) represented the minimum volume the user could hear at a given frequency. This is a device relative volume. Zero represents no sound, so the minimum possible value is one (1), and the maximum (on the device tested) is fifteen (15). As

such, higher values in this column represent worse scores on hearing evaluation. Unheard frequencies are marked with an “X”

• Column 8 (Speech Disc. Score), represent the scores obtained on the Speech Discrimination exam. The score is present as the percentage of correct answers given after 20 trials.

• Column 9 indicates whether the user reported bad hearing or not. This is normally either “Yes” or “No”, however, user 0006 indicated that while he didn't notice (other) audition problems, he did feel that he had difficulty understanding words.

Normal Hearing Reported

First, it was attempted to establish a baseline for what the average results would be for a normal hearing subject. So, Table 4 was filtered to include only users that didn't report hearing problems. It's important to remember that the volunteer with code 0006 reported a sense of difficulty understanding people. It can be seen that his scores do in fact stand out from the users which reported normal hearing. When compared against normal hearing subjects, all his volume scores are among the highest of each frequency. Additionally, one should look at his Speech Discrimination score: among the users with “normal” hearing, he was the only one below 95%, in agreement with his perception of difficulty understanding people. Both these situations seem to indicate that he has in fact subpar audition and made us to include him in the same group of those which reported problematic hearing.

As such, the following table was created:

Table 5. Improved Normal Hearing Audition Test Results

Code	250	500	1000	2000	4000	8000	Speech Disc. Score	Problem
0002 – right	2	1	1	1	3	4	95%	No
0002 – left	1	1	1	1	1	6		
0004 – right	2	1	2	1	1	2	100%	No
0004 – left	2	1	1	1	1	1		
0005 – right	2	1	1	1	1	1	95%	No
0005 – left	1	1	1	1	1	1		
0009 – right	1	1	1	1	1	1	95%	No
0009 – left	1	1	1	1	1	2		
0010 – right	1	1	2	1	1	3	100%	No
0010 – left	1	1	1	1	1	4		

Here it can be seen that the Speech Discrimination scores are all between 95 and 100%, meaning the user got at most one incorrect answer. The PTA scores are all much more balanced, all values ranging from 1 to 2 except for the highest frequencies. By averaging these scores a graph of the expected results from this exam, in normal hearing users, was drawn. This generated the blue line that can be seen in Figure 45, where the standard deviation can also be seen.

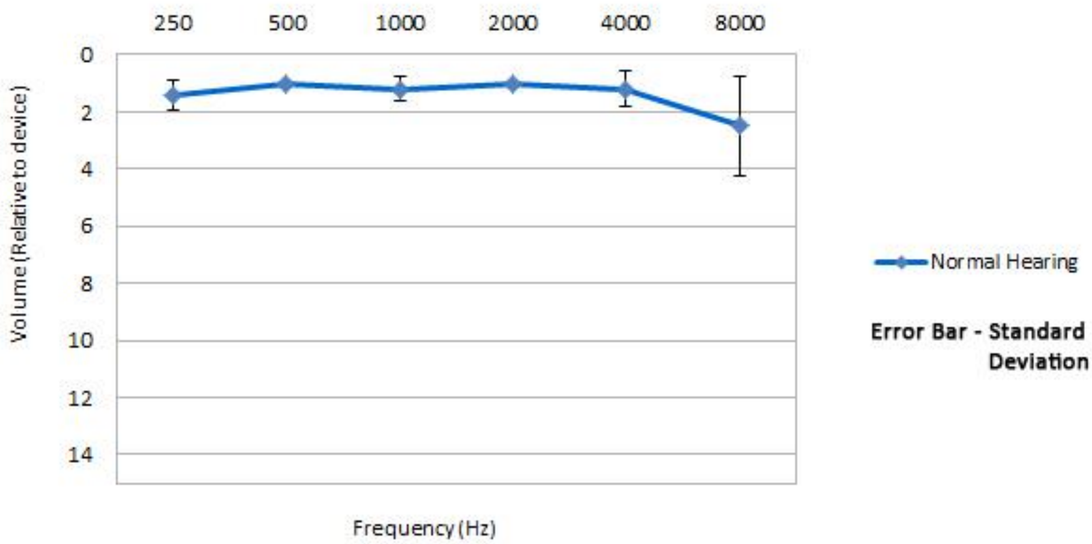

Figure 45. Normal Hearing Subjects Average

Faulty Hearing Reported

A similar procedure was followed in order to create the table which included only users which reported faulty hearing, visible in Table 6. Each user's exam result will now be analyzed in an attempt to identify a correlation between poor exam results and faulty hearing.

Table 6. Faulty Hearing Audition Test Results

Code	250	500	1000	2000	4000	8000	Speech Disc. Score	Problem
0003 - right	3	1	1	1	2	6	95%	Yes
0003 - left	1	1	1	1	3	8		
0006 - right	2	1	2	1	5	3	85%	Maybe
0006 - left	1	1	1	2	3	3		
0007 – right (with aid)	1	1	1	3	2	6	95%	Yes
0007 - left	7	7	6	8	X	7		
0007 – right (without aid)	1	1	1	7	8	8		
0008 - right	5	4	3	2	7	X	80%	Yes
0008 - left	1	1	1	5	5	X		
0011 - right	2	1	1	3	6	X	85 %	Yes
0011 - left	2	1	1	3	4	X		

From this table, it is visible that their Speech Discrimination scores are unlike the normal hearing subjects. None of these users had a Speech Discrimination score of 100%, and only two of them achieved 95%. If we take into account that one of these two users performed the exam using a hearing aid (0007), that number could possibly be even smaller. This is a good sign that this might be an indicator of problematic hearing.

To further evaluate the results, each abnormal hearing subjects score will now be compared with the average obtained earlier from normal hearing volunteers. From the data presented in Table 6 an audiogram for each subject will be featured, comparing his left and right ear to the baseline obtained from the Normal Hearing subjects.

User 0003

This user reported bad audition, but no reference to a hearing aid was made. From comparing his audiogram with the average audiogram (see Figure 46), it is visible that in both ears he requires a high elevation of volume for the 4000 and 8000 Hz trials, as well as a small increase on the 250 Hz frequency for the right ear.

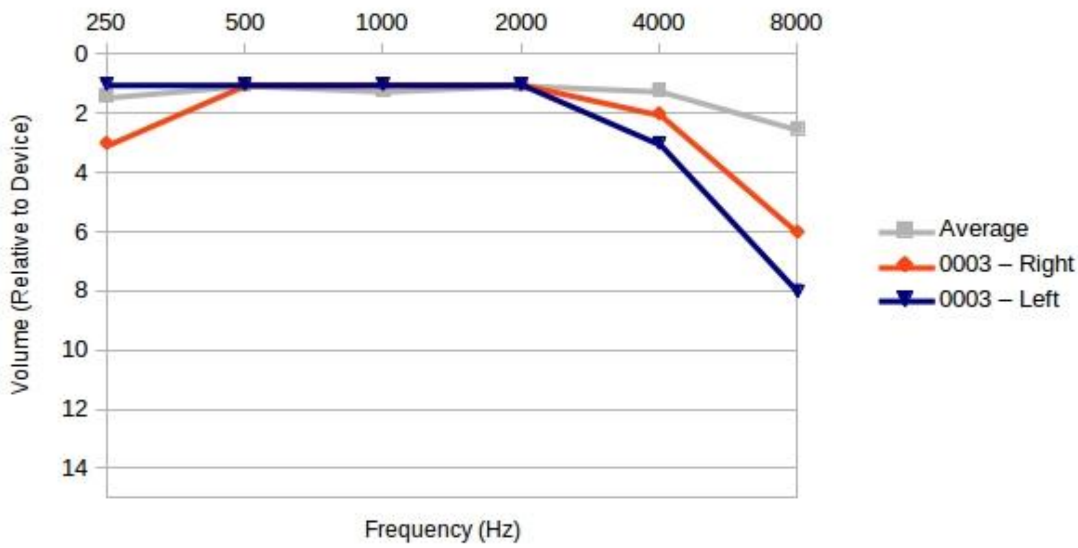

Figure 46. Audiogram Results (User Code 0003)

These problematic frequencies might be the cause of the reported audition troubles.

User 0006

This is the user that reported good hearing, while at the same time having a sense of having difficulty in understanding speech. His audiogram (shown in Figure 47), while not requiring very high volumes, shows a slight variance when compared to the average audiogram.

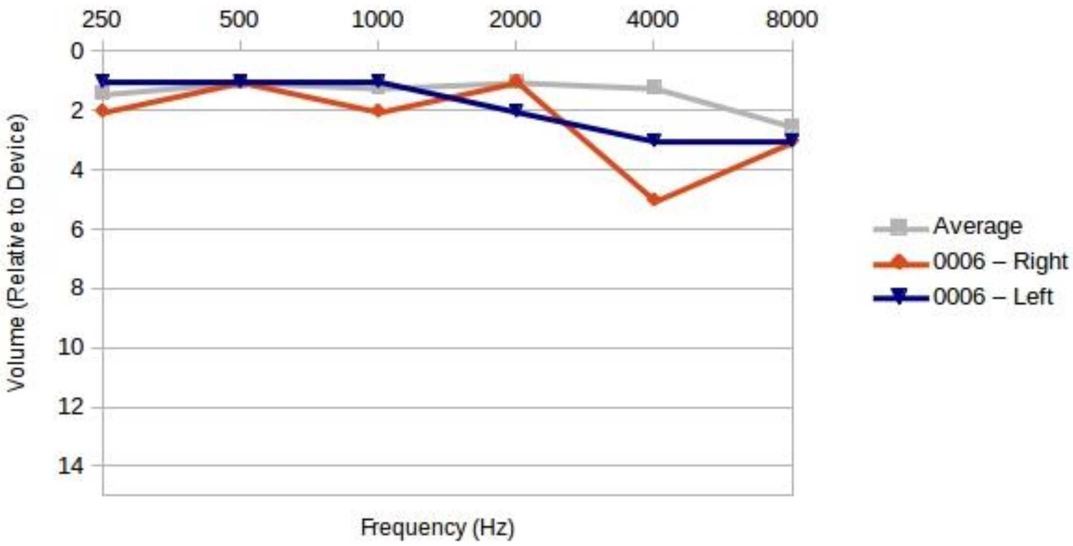

Figure 47. Audiogram Results (User Code 0006)

If we take into consideration these deviations and the fact that his Speech Discrimination score (85%) was below the values obtained by those with normal hearing, one might be able to understand his reported hearing difficulties.

User 0007

User 0007 is an important subject when analyzing the exams results obtained from prototype based testing. This user reported having very difficult hearing: he had one extremely bad ear, his left one, which was seen as a non-factor for his hearing; his right ear requires the use of a hearing aid.

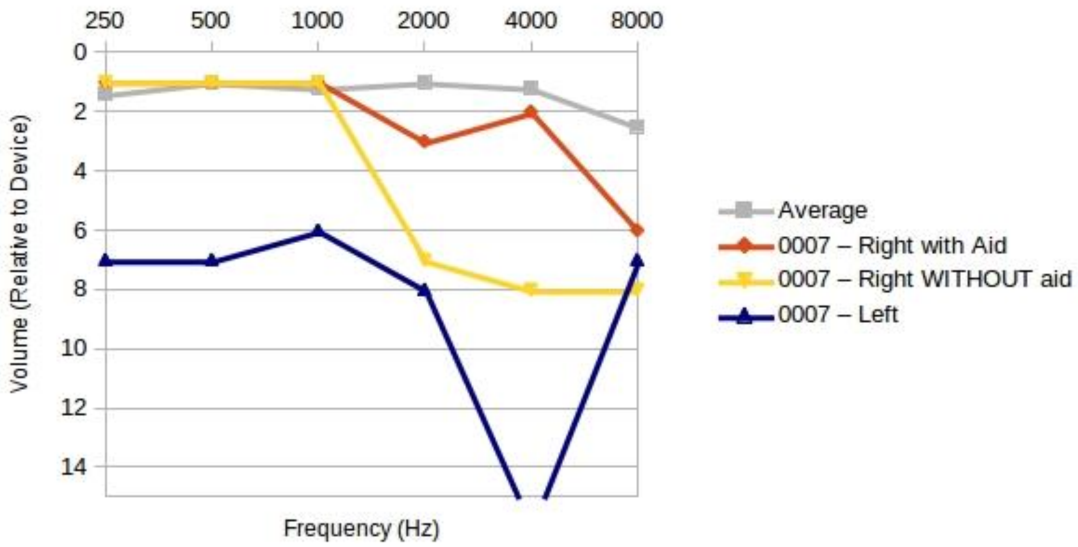

Figure 48. Audiogram Results (User Code 0007)

By analyzing Figure 48, his audiograms can be compared against the average. The one with the color red, represents his right ear when using the hearing aid. It is still visible that his hearing is flawed, ranking well below average and worse than the previous faulty hearing users. However, when compared with the unaided audiogram (yellow line), it is possible to see how much better his hearing is when he uses the hearing aid. This observable drop in performance is important in showing that the system is capable of detecting this type of variations in the user.

Lastly, his left hear (blue line), which was reported as a non-factor, show scores that ranked well below any other taken in this test. From 250 to 2000 Hz he required volumes unseen before and was even completely unable to identify sounds at 4000 Hz. This is represented by the line going beyond the volume limit of 15.

This user appears to be a valuable test subject, because he provides a nice array of different samples, exemplifying the variance of results that the system might be able to detect in real life.

User 0008

User 0008 reported problematic hearing and that he was professionally instructed to use a hearing aid. Even though he owned a hearing aid, he did not use it in his daily life because he felt uncomfortable doing so. Because of this, these audiograms reflect his performance without hearing aid, therefore, sub-par results are expected.

By analyzing Figure 49, it is visible that both his ears yielded bad results. His left ear scores the normal values at lower frequencies, but increasingly worsens after 1000 Hz, being completely unable to hear the sounds at 8000 Hz. Even worse is his right ear, which at all frequencies is below average, and shows further difficulty hearing sounds at 4000 Hz and 8000 Hz. One thing to notice is the relatively good score of the right ear at 2000 Hz. This is a case where a retest (of the frequency or audiogram) should be made to mitigate possible unwanted and incorrect answers that might have been given to the system by the user.

Further reinforcing the conclusions that this user does in fact face some problems with his audition, and that a hearing aid is recommended, is the fact that he scored the lowest of all results in Speech Discrimination, obtaining only 80% of correct answers.

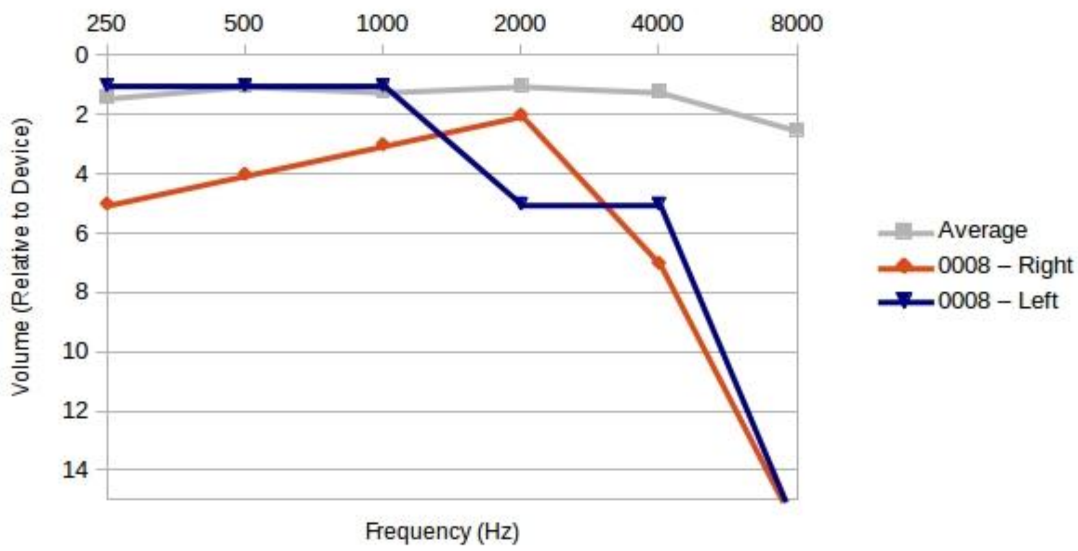

Figure 49. Audiogram Results (User Code 0008)

This user is also a valuable test subject, since his unwillingness to use the recommended hearing aid provided us the bad results that are to be expected and that show an ability to detect such problems by using the system developed.

User 0011

Our last test subject also reported faulty hearing. His audiogram, in Figure 50, shows increasingly more difficulty in listening to sounds after the 1000 Hz mark, rendering him completely unable to hear sounds at 8000 Hz. His score in the Speech Discrimination exam shows an equally bad result, answering correctly only 85% of the times.

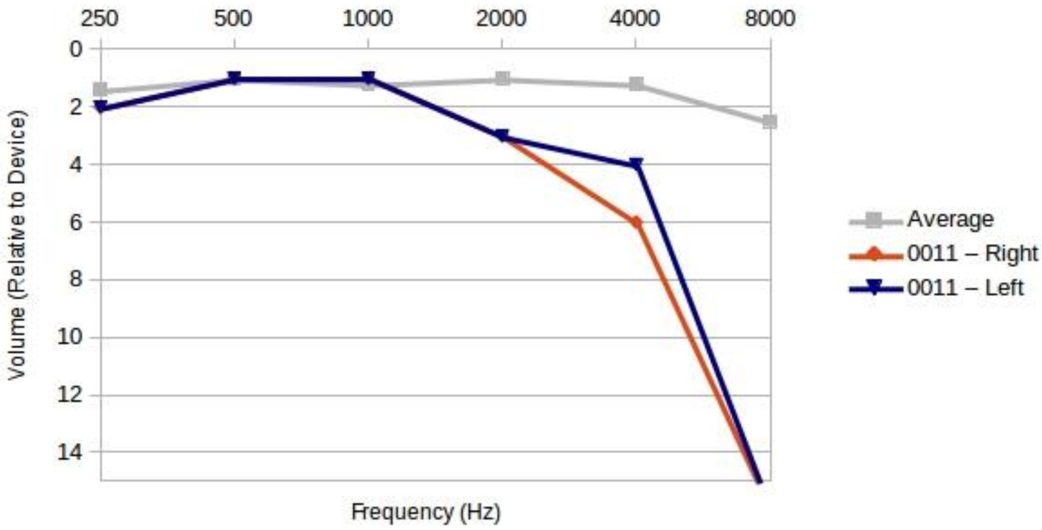

Figure 50. Audiogram Results (User Code 0011)

Again, the bad hearing reported and the expected bad results obtained on both exams, lead us to believe that our system is capable of detecting flaws on the users hearing system, which is valuable in the validation of these tests.

4.6.3 System Usability Scale Questionnaire

After completing tasks, the volunteers were given a System Usability Scale (SUS) questionnaire in order to express their opinion on the system. They were also given two open ended questions regarding the prototype. The entire Questionnaire can be found in Appendix A- SUS Questionnaire. This section presents the data gathered from the results obtained.

In a SUS Questionnaire, users are presented with a Likert Scale comprising 10 questions. Each question has a grading system from 1 to 5, where 1 is "Completely disagree" and 5 is "Completely agree", and the user should select the option he feels most adequately represents his feelings when using the prototype. Due to the way this questionnaire is structured, odd numbered questions represent lines where it is hoped that the user will give a high score (4 or 5) and the even numbered questions represent lines where it is hoped that the user will give a low score (1 or 2).

Table 7 shows the scores obtained by the test subjects.

Table 7. SUS Questionnaire Results

User Code	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
0002	5	5	5	1	5	1	5	5	5	1
0003	5	1	5	3	5	2	5	2	5	2
0004	3	3	4	4	5	4	2	2	5	2
0005	5	1	5	3	5	1	3	1	5	3
0006	5	3	5	5	5	3	4	1	5	2
0007	4	4	4	4	4	1	4	1	4	4
0008	5	1	5	1	5	1	5	1	5	1
0009	5	1	5	1	5	1	5	1	5	2
0010	5	1	5	5	5	1	3	1	5	4
0011	5	1	5	1	5	1	3	1	5	1

To obtain the final value, first we have to go through the following steps (for each user):

- Scores from odd questions (Q1, Q3, Q5, Q7 and Q9) have their value decreased by one. For example: if Q1 is given a score of 4 by the user, that line will be worth 3 points (4 minus 1). The same procedure is followed for each odd numbered question.
- Scores from even questions (Q2, Q4, Q6, Q8 and Q10) are subtracted from the value 5. For example: if Q2 is given a score of 1, that line will be worth 4 points (5 minus 1). The same procedure is followed for each even numbered question.
- The scores obtained in the previous steps are summed up and, finally, that total is multiplied by 2.5 (to get a score between 0 and 100).

The resulting scores for each user can be seen below:

Table 8. SUS Questionnaire Scores

User Code	Score
0002	80
0003	87,5
0004	60
0005	85
0006	75
0007	65
0008	100
0009	97,5
0010	77,5
0011	95
Mean	82,25
Standard Deviation	13.40

Research indicate that the average value of SUS questionnaire is defined as 68 [93]–[95]. So the mean score of 82.25 obtained by all users of the system seems to point to a good evaluation. All but two users ranked the system higher than 68. From these two, user with code 0004 (which is also the one with minimum value) might be explained by the fact that he was one of the users with longest test time, still being tested at all tasks and, due to a technical problem, he was also forced to repeat the distance visual test multiple times. This seems to be in agreement with his answer to "What was the biggest challenged you faced in the system?", which was "The maneuvering of the device". The other user which is just barely below average was user 0007 (65 vs 68), but in this case, he answered to the same question: "I did well with the system". He had no need to repeat any exam, so it is harder to justify this score. It might be simply due to a misunderstanding of some questions, or maybe it is truly an indicator that some things need to be improved (which, unfortunately, were not identified concretely by this or other users).

In summary, the results obtained from the users seem to indicate that system was well received and that the users would feel comfortable using it.

4.7 Summary

The tests performed allowed us to evaluate three components of our exams:

- Vision Exams – All vision exams were performed. A strong correlation between paper and prototype results was found in Distance Visual acuity. This is seen as a strong and valuable feature of this system, because such an important exam is normally hard to reproduce accurately at the user's home. Near vision acuity, also produced some hints that it could help to evaluate and track the progress of the user's eyesight. However, it was ultimately too inaccurate and the procedure raised too many questions on its validity to be considered as a definite tool. Therefore this feature requires further testing to draw stronger conclusions. Contrast Sensitivity and Shape Discrimination lacked the traditional tests so that a result comparison could be made, and did not produce strong relations when compared to the Distance visual acuity or Near vision results. Additionally, a (probably unconscious) tendency to reduce testing distance was observed, which leads us to believe that the distance measure tool, after being validated and integrated with the vision exam (by being constantly active instead of just prior to the exam) might be an important asset in helping achieve accurate results.

- Audition exams – The work performed with the ten testing subjects showed a good relation between the users who reported problems and poor performance on our exams. The only subjects that scored less than 95% on Speech Discrimination score were the users who reported faulty hearing. From those with faulty hearing, only two scored 95% and one of them did so while using his hearing aid, which may have positively influenced his performance. Through our analysis of the PTA exams it was also possible to see that users with problematic hearing did indeed perform poorly when compared to normal hearing subjects. The audiograms produced by the problematic users was indeed worse than normal hearing subject, and that was easily observable by the volume they required to be able to hear a sound and in the fact that sounds at some frequencies were sometimes completely inaudible. This seems to indicate a good correlation between bad hearing and bad performance on the implemented audition exams.

- Usability – The users were able to perform all the exams, and showed an increase in comfort when repeating the same task. The questionnaire reinforced the notion that the user would be capable and willing of operating this system on his own. This is an important factor, while a certain difficulty is to be expected at first from unexperienced subjects. It is important that they feel comfortable using the system at their homes, without the need for assistance by of another person. A growing comfort level is expected after the first times taking the exams, since the user better understands how to interact with the system and avoids making mistakes.

Still, there is room for improvements. While a strong correlation of one of the vision exams was visible, further testing would be required to validate the remaining games (and even further validate the distance visual acuity one). Since our near vision game did not match exactly the procedure of the standard test, it would be valuable to change its procedure or replicate this exam in a chart, card or similar form, properly validated, and then compare the results. On the audition testing section, a calibration of volume output and a validation against clinically conducted trials would be an incredibly increase in test reliability and precision. This would allow us to move from a device relative volume system to a measurable volume unit that could be compared between devices and against a clinically performed PTA exam. Additionally, a change could be made to the PTA exam procedure, instead of the volume increase being linear it could perhaps be dependent on previous answers; this would help decrease the relatively high testing time of abnormal ears (when compared to normal hearing). Another change to reinforce reliability could be the retest of all or just the dubious answers before committing to defining the minimum volume of a given frequency. However this could lead to longer test times, which is not desired, or have the opposite effect and decrease reliability, by exhausting the user and therefore having him less focus on the stimuli. Lastly, some potential interface and test procedure changes were identified through volunteer observation during testing. These changes could possibly lower the complexity of the system and further entice the users to have their senses tested on a regular basis.

Chapter 5

Final Conclusions

Taking into account the growth in elderly population and the risks that they face in their daily lives it becomes more and more important to provide these citizens with the assistance they require. Due to the increased challenges that these people face, leading to a loss of independence, it becomes imperative to give these users the means to keep constant surveillance on themselves.

Smartphones are powerful devices that are considerable easy to acquire and manage. If we can combine the availability and advantages of these devices and provide the users with the tools they require to monitor themselves, disastrous situations could possibly be prevented.

One such area that is worth the elderlies' attention is sense monitoring, namely vision and audition. These two senses represent a significant part of the information channels which people use to obtain information in their daily lives. Yet, the tests currently available that allow the evaluation of these senses have a long period of time between testing sessions, leading to a large time gap where the user is left untested, even despite the fact that changes can be sudden. The sooner one can detect the problem, the sooner precautions can be taken and disasters can be avoided. Therefore, it becomes important to provide these citizens with ways that permit them to evaluate their senses, while waiting for their usual clinical visit, so that if a problem is detected, a more urgent appointment can be set up. All these factors lead us to the goal of this dissertation: the design, implementation and evaluation of methodologies which allow the users to evaluate their vision and audition, without the need of external help, in the time intervals where they have to wait for a clinical exam. The point of this system would not be to replace or delay a clinical visit, but to provide the user with a reliable everyday tool that they can use to track their progress, in an attempt to find potential sudden or progressive changes that would be detected too late otherwise.

Research on current similar systems, which offer the user with ways to track some parts of their vision or audition, was conducted. These systems suffer from many shortcomings that this dissertation attempts to correct, such as: the need for external help of a care-taker or professional, the focus on just one sense or just a part of a sense.

In the work developed, multiple characteristics of each sense were targeted, hoping to evaluate a large range of factors that might impact the user’s everyday life. Methodologies that allow users to conduct exams on themselves, without the need of external help or prior training were developed. Additionally, they were designed to require few extra accessories in order to keep the costs of the user to a minimum, thus being available to more users. As such, the system provides the user with a total of 6 exams, two of them for Audition evaluation (Pure Tone Audiometry and Speech Discrimination) and four for Vision evaluation (Distance Visual Acuity, Near Vision, Contrast Sensitivity and Shape Discrimination). To perform all the exams available a user only requires (besides the obvious smartphone) a regular headphone set (cheap and easily acquirable) and any item that will serve as an eye patch.

Trials on a sample of volunteers were performed. Vision and Audition related data was gathered from these volunteers and paper-based and prototype-based tests were conducted. The comparison of these results allowed for an evaluation of the system:

- The distance visual acuity exam, one of the most important vision tests, yielded a good correlation between paper and prototype based testing. This test is usually hard to reproduce, due to a variety of reasons, such as: distance required, help from another person and chart variability (to avoid memorization). Allowing the user to track this sense at home is thus very important.
- The near vision reading exam was not as conclusive. When compared against the paper-based reading test it achieved similar, but constantly lower results. This might be related to the fact that the testing process is not exactly the same. In the paper test the user only has to read the sentence, but in the prototype test he has to complete it (which led to a higher difficulty than initially thought). In its current state, the methodology used might be useful to track changes, but without being able to pinpoint the actual value, the methodology used should be re-examined. When compared against the paper-based letter acuity test the results were highly variable, with some users obtaining similar results, but others obtaining very different values. This test is even more different in testing process and in the type of result obtained (visual acuity vs reading acuity), thus it is even harder to obtain a correlation between results. The results obtained from comparing these tests lead us to believe that further validation is required or even a test restructuring, either by changing the test methodology or by creating and validating a paper based test that reproduces the exam procedure.
- Audition exams results showed a good correlation between users that indicated having problematic hearing and bad results. Both the pure tone audiometry exam and the speech discrimination exam obtained worse results when performed by users with problematic hearing. Left and right ear differentiation also showed the expected results when the user identified the side which was most problematic. Additionally, there was a user who reported not yet being diagnosed with a problem, but that thought he had difficulty understanding speech. This user scores seemed to point to the fact that he, indeed, had troubles, due to his scores being significantly worse than normal hearing subjects.

- Contrast Sensitivity and Shape Discrimination results were also obtained; however, due to the lack of their traditional test results, it was not possible to compare the results obtained in the exams and draw conclusions about their accuracy. Additionally, a correlation between the scores obtained in those tests and in distance and near vision testing was not found. As such, in order to validate the results from these exams, the traditional tests need to be performed.

These results seem to indicate that the methodologies developed can be seen as a tool that the user can take advantage of on his everyday life, which will allow him to track variations on his vision or audition quality.

One big challenge ahead is the further validation of the exams against clinical results. While a good correlation was found between some test results and qualitative information provided by the users, if the values obtained from our exams were compared against concrete clinical tests, this could help further strengthen the reliability of these tests.

Overall, the goals set for this dissertation were achieved. Methodologies which allow for an easy to use testing procedure and that were well received (by the testing sample) were created. The results proved to be useful in the evaluation of vision and audition quality. The work developed was self-contained, cheap and usable by a single person, meaning that it was able to solve some of the limitations found on the available tools found during the research step.

5.1 Future Work

Regarding the prototype system developed, room for improvement has been identified. During testing, potential interface changes that could use a makeover, in order to be more intuitive to the user, were identified. Also during test sessions, some potential need for alterations were found: one of this was the Distance Visual Acuity exam procedure, where the rotation of the device, while workable, presented some challenges to this senior population sample due to their difficulty of managing the device without pressing the device specific buttons; another significant point, was the need to fight the inherent instinct to reduce the eye-device distance after a vision game is initiated. The former might be solvable with a simple change to the expected movement like reported in the Result Evaluation chapter. The latter could perhaps be solved or mitigated by improving the current distance measure tool and integrating it with the exams, so that it is running on the background and can alert the user to irregular test distances.

Additionally, to give the exams higher credibility, conducting trials in a clinical setting with professionals using their tools and expertise would be of utmost importance. This would provide us with a more quantitative measure of our results. This would also allow and require a calibration of our Pure Tone Audiometry xexam, so that each volume outputted by the devices (in combination with a defined headphone set) could be measured, instead of using the current device relative system.

Finally, taking into account this dissertation's integration in Fraunhofer Portugal's project of fall risk assessment through the evaluation of different metrics, a relation between deficiencies in any or both of the senses involved here and the fall risk of a given user could be analyzed. If a relation was found, it might prove useful in helping to prevent one of the most critical risk the elderly population faces today (falls).

References

[1] S. R. LORD, R. D. CLARK, and I. W. WEBSTER, “Visual Acuity and Contrast Sensitivity in Relation to Falls in an Elderly Population,” *Age Ageing*, vol. 20, no. 3, pp. 175–181, May 1991.

[2] S. R. Lord and J. Dayhew, “Visual risk factors for falls in older people.,” *J. Am. Geriatr. Soc.*, vol. 49, no. 5, pp. 508–15, May 2001.

[3] E. L. Lamoureux, E. L. Lamoreux, E. Chong, J. J. Wang, S. M. Saw, T. Aung, P. Mitchell, and T. Y. Wong, “Visual impairment, causes of vision loss, and falls: the singapore malay eye study.,” *Invest. Ophthalmol. Vis. Sci.*, vol. 49, no. 2, pp. 528–33, Feb. 2008.

[4] A. P. Chhetri, F. Wen, Y. Wang, and K. Zhang, “Shape Discrimination Test on Handheld Devices for Patient Self-Test,” pp. 502–506, 2010.

[5] I. a Falkenstein, D. E. Cochran, S. P. Azen, L. Dustin, A. M. Tammewar, I. Kozak, and W. R. Freeman, “Comparison of visual acuity in macular degeneration patients measured with snellen and early treatment diabetic retinopathy study charts.,” *Ophthalmology*, vol. 115, no. 2, pp. 319–23, Feb. 2008.

[6] a T. Johnson, C. R. Dooly, and C. R. Simpson, “Generating the Snellen Chart by computer.,” *Comput. Methods Programs Biomed.*, vol. 57, no. 3, pp. 161–6, Nov. 1998.

[7] T. W. Raasch, I. L. Bailey, and M. A. Bullimore, “Repeatability of visual acuity measurement.,” *Optom. Vis. Sci.*, vol. 75, no. 5, pp. 342–8, May 1998.

[8] D. A. Rosser, D. A. Laidlaw, and I. E. Murdoch, “The development of a ‘reduced logMAR’ visual acuity chart for use in routine clinical practice.,” *Br. J. Ophthalmol.*, vol. 85, no. 4, pp. 432–6, Apr. 2001.

[9] M. Camparini, P. Cassinari, L. Ferrigno, and C. Macaluso, “ETDRS-Fast: Implementing Psychophysical Adaptive Methods to Standardized Visual Acuity Measurement with ETDRS Charts,” *Invest. Ophthalmol. Vis. Sci.*, vol. 42, no. 6, pp. 1226–1231, May 2001.

[10] “Calculate Size of 20/20 Letter.” [Online]. Available: http://www.blackwelleyesight.com/eye-math/2020-letter/.

[11] P. K. Kaiser, “Prospective evaluation of visual acuity assessment: a comparison of snellen versus ETDRS charts in clinical practice (An AOS Thesis).,” *Trans. Am. Ophthalmol. Soc.*, vol. 107, pp. 311–24, Dec. 2009.

[12] R. J. Dennis, J. M. a Beer, J. B. Baldwin, D. J. Ivan, F. J. Lorusso, and W. T. Thompson, “Using the Freiburg Acuity and Contrast Test to measure visual performance in USAF personnel after PRK.,” *Optom. Vis. Sci.*, vol. 81, no. 7, pp. 516–24, Jul. 2004.

[13] A. Colenbrander and S. Francisco, “Measuring Vision and Vision Loss,” vol. 5, pp. 1–42, 2001.

[14] C. Kniestedt and R. L. Stamper, “Visual acuity and its measurement.,” *Ophthalmol. Clin. North Am.*, vol. 16, no. 2, pp. 155–70, v, Jun. 2003.

[15] W. Wesemann, “[Visual acuity measured via the Freiburg visual acuity test (FVT), Bailey Lovie chart and Landolt Ring chart].,” *Klin. Monbl. Augenheilkd.*, vol. 219, no. 9, pp. 660–7, Sep. 2002.

[16] M. Ma, “Normal values for the Pelli-Robson,” vol. 3350, no. 00, 2001.

[17] D. G. Pelli and P. Bex, “Measuring contrast sensitivity.,” *Vision Res.*, vol. 90, pp. 10–4, Sep. 2013.

[18] M. Bach, “The Freiburg Visual Acuity test--automatic measurement of visual acuity.,” *Optometry and vision science : official publication of the American Academy of Optometry*, vol. 73, no. 1. pp. 49–53, Jan-1996.

[19] M. Bach, “The Freiburg Visual Acuity Test-variability unchanged by post-hoc re-analysis.,” *Graefes Arch. Clin. Exp. Ophthalmol.*, vol. 245, no. 7, pp. 965–71, Jul. 2007.

[20] A. S. Hawkins, J. P. Szlyk, Z. Ardickas, K. R. Alexander, and J. T. Wilensky, “Comparison of contrast sensitivity, visual acuity, and Humphrey visual field testing in patients with glaucoma.,” *J. Glaucoma*, vol. 12, no. 2, pp. 134–8, Apr. 2003.

[21] K. Adigun, T. S. Oluleye, M. M. Ladipo, and S. A. Olowookere, “Quality of life in patients with visual impairment in Ibadan: a clinical study in primary care.,” *J. Multidiscip. Healthc.*, vol. 7, pp. 173–8, Jan. 2014.

[22] A. De Bortoli and O. Gaggi, “PlayWithEyes: A new way to test children eyes,” *2011 IEEE 1st Int. Conf. Serious Games Appl. Heal.*, pp. 1–4, 2011.

[23] L. Cyert, V. Dobson, M. T. Kulp, M. G. Maguire, G. Qunin, P. P. Schmidt, B. Moore, E. Peskin, E. Ciner, D. Orel-Bixler, M. Redford, and J. Schultz, “Preschool visual acuity screening with HOTV and Lea symbols: testability and between-test agreement,” vol. 81, no. 9, pp. 678 – 83.

[24] M. E. Mercer, J. R. Drover, K. J. Penney, M. L. Courage, and R. J. Adams, “Comparison of Patti Pics and Lea Symbols optotypes in children and adults.,” *Optom. Vis. Sci.*, vol. 90, no. 3, pp. 236–41, Mar. 2013.

[25] K. R. ALEXANDER, W. XIE, and D. J. DERLACKI, “Visual Acuity and Contrast Sensitivity for Individual Sloan Letters,” *Vision Res.*, vol. 37, no. 6, pp. 813–819, Mar. 1997.

[26] “5 Reasons To Use Sloan Letters.” [Online]. Available: http://blog.schoolhealth.com/bid/58533/Selecting-an-Eye-Chart-Sloan-Letters-or-Snellen.

[27] M. Bach, “Freiburg Vision Test.”

[28] A. Pentland, “Notes and Comment Maximum likelihood estimation : The best PEST,” vol. 28, no. 4, pp. 377–379, 1980.

[29] H. R. Lieberman and A. P. Pentland, “Microcomputer-based estimation of psychophysical thresholds: The Best PEST,” *Behav. Res. Methods Instrum.*, vol. 14, no. 1, pp. 21–25, Jan. 1982.

[30] D. Thomson, “VA testing in optometric practice Part 2 : Newer chart designs,” no. m, pp. 22–24, 2005.

[31] G. J. J. and I. E. M. R R A Bourne, D A Rosser, P Sukudom, B Dineen, D A H Laidlaw, “Evaluating a new logMAR chart designed to improve visual acuity assessment in population-based surveys.” [Online]. Available: http://www.nature.com/eye/journal/v17/n6/fig_tab/6700500t2.html.

[32] “Logmar Definition & Protocol.” [Online]. Available: http://www.sussexvision.co.uk/info_pages.php?pages_id=17.

[33] R.-M. M. Oduntan OA, Mashige KP, “A comparison of two methods of logMAR visual acuity data scoring for statistical analysis,” *S Afr Optom 68*, pp. 155–163, 2009.

[34] “LogMAR Test.” [Online]. Available: http://www.sussexvision.co.uk/info_pages.php?pages_id=17.

[35] A. K. Bittner, P. Jeter, and G. Dagnelie, “Grating acuity and contrast tests for clinical trials of severe vision loss.,” *Optom. Vis. Sci.*, vol. 88, no. 10, pp. 1153–63, Oct. 2011.

[36] J. M. Wood, P. F. Lacherez, A. a Black, M. H. Cole, M. Y. Boon, and G. K. Kerr, “Postural stability and gait among older adults with age-related maculopathy.,” *Invest. Ophthalmol. Vis. Sci.*, vol. 50, no. 1, pp. 482–7, Jan. 2009.

[37] “Jaeger Eye Chart.” [Online]. Available: http://www.allaboutvision.com/eye-test/.

[38] A. Colenbrander and S. Francisco, “Measuring Vision and Vision Loss,” vol. 5, pp. 1–42, 2001.

[39] L. L. Sloan, “Needs for precise measures of acuity. Equipment to meet these needs.,” *Arch. Ophthalmol.*, vol. 98, no. 2, pp. 286–90, Feb. 1980.

[40] G. A. Hahn, D. Penka, C. Gehrlich, A. Messias, M. Weismann, L. Hyvärinen, M. Leinonen, M. Feely, G. Rubin, C. Dauxerre, F. Vital-Durand, S. Featherston, K. Dietz, and S. Trauzettel-Klosinski, “New standardised texts for assessing reading performance in four European languages.,” *Br. J. Ophthalmol.*, vol. 90, no. 4, pp. 480–4, Apr. 2006.

[41] M. C. Burggraaff, R. M. A. van Nispen, S. Hoek, D. L. Knol, and G. H. M. B. van Rens, "Feasibility of the Radner Reading Charts in low-vision patients.," *Graefes Arch. Clin. Exp. Ophthalmol.*, vol. 248, no. 11, pp. 1631–7, Nov. 2010.

[42] Y. Z. Wang, E. Wilson, K. G. Locke, and a O. Edwards, "Shape discrimination in age-related macular degeneration," *Invest Ophthalmol Vis Sci*, vol. 43, no. 6, pp. 2055–2062, 2002.

[43] R. A. Schuchard, "Validity and Interpretation of Amsler Grid Reports," *Arch. Ophthalmol.*, vol. 111, no. 6, p. 776, Jun. 1993.

[44] J. Haugsdal and S. Elliott, "Age-Related Macular Degeneration: From One Medical Student to Another," 2013. [Online]. Available: http://webeye.ophth.uiowa.edu/eyeforum/tutorials/AMD-medical-student/.

[45] D. C. Fletcher, R. A. Schuchard, and L. W. Renninger, "Patient awareness of binocular central scotoma in age-related macular degeneration.," *Optom. Vis. Sci.*, vol. 89, no. 9, pp. 1395–8, Sep. 2012.

[46] W. A. Fletcher, "Acute Idiopathic Blind Spot Enlargement," *Arch. Ophthalmol.*, vol. 106, no. 1, p. 44, Jan. 1988.

[47] Y.-Z. Wang, Y.-G. He, G. Mitzel, S. Zhang, and M. Bartlett, "Handheld shape discrimination hyperacuity test on a mobile device for remote monitoring of visual function in maculopathy.," *Invest. Ophthalmol. Vis. Sci.*, vol. 54, no. 8, pp. 5497–505, Aug. 2013.

[48] G. Cummings, "Vision screening in junior schools," *Public Health*, vol. 110, no. 6, pp. 369–372, Nov. 1996.

[49] J. Birch, "Identification of red-green colour deficiency: sensitivity of the Ishihara and American Optical Company (Hard, Rand and Rittler) pseudo-isochromatic plates to identify slight anomalous trichromatism.," *Ophthalmic Physiol. Opt.*, vol. 30, no. 5, pp. 667–71, Sep. 2010.

[50] "Understanding your Audiogram." [Online]. Available: http://www.hopkinsmedicine.org/hearing/hearing_testing/understanding_audiogram.html .

[51] S. A. Gelfand, "Use of the carrier phrase in live voice speech discrimination testing."

[52] A. Y. Bonino, L. J. Leibold, and E. Buss, "Release from perceptual masking for children and adults: benefit of a carrier phrase.," *Ear Hear.*, vol. 34, no. 1, pp. 3–14.

[53] R. Roeser and M. Valente, *AUDIOLOGY, 3-Volume Set: Diagnosis, Treatment and Practice Management*. Thieme, 2007, p. 2000.

[54] K. Ullrich and D. Grimm, "Most comfortable listening level presentation versus maximum discrimination for word discrimination material.," *Audiology*, vol. 15, no. 4, pp. 338–47.

[55] R. G. W. Jerry J Bennett, “Relationships Between Speech Discrimination Ability and Comfortable Listening Levels.” .

[56] “Hearing in Noise Test (HINT).” [Online]. Available: http://www.californiaearinstitute.com/audiology-services-hint-bay-area-ca.php.

[57] J. Chhablani, S. Kaja, and V. A. Shah, “Smartphones in ophthalmology.,” *Indian J. Ophthalmol.*, vol. 60, no. 2, pp. 127–31.

[58] B. V Stanzel and C. H. Meyer, “[Smartphones in ophthalmology : Relief or toys for physicians?].,” *Ophthalmologe*, vol. 109, no. 1, pp. 8–20, Jan. 2012.

[59] J. M. Black, R. J. Jacobs, G. Phillips, L. Chen, E. Tan, a. Tran, and B. Thompson, “An assessment of the iPad as a testing platform for distance visual acuity in adults,” *BMJ Open*, vol. 3, no. 6, pp. e002730–e002730, Jan. 2013.

[60] Y. Z. Wang, “Effects of aging on shape discrimination.,” *Optom. Vis. Sci.*, vol. 78, no. 6, pp. 447–54, Jun. 2001.

[61] “Peek Vision.”

[62] L. J. Haddock, D. Y. Kim, and S. Mukai, “Simple, inexpensive technique for high-quality smartphone fundus photography in human and animal eyes.,” *J. Ophthalmol.*, vol. 2013, p. 518479, Jan. 2013.

[63] A. C. S. Kam, J. K. K. Sung, T. Lee, T. K. C. Wong, and A. van Hasselt, “Clinical evaluation of a computerized self-administered hearing test,” *Int. J. Audiol.*, vol. 51, no. 8, pp. 606–610, Aug. 2012.

[64] J. M. Choi, J. Sohn, Y. Ku, D. Kim, and J. Lee, “Phoneme-Based Self Hearing Assessment on a Smartphone,” vol. 17, no. 3, pp. 526–529, 2013.

[65] C. Stephanidis and M. Antona, Eds., *Universal Access in Human-Computer Interaction. Aging and Assistive Environments*, vol. 8515. Cham: Springer International Publishing, 2014.

[66] H. Hwangbo, S. H. Yoon, B. S. Jin, Y. S. Han, and Y. G. Ji, “A Study of Pointing Performance of Elderly Users on Smartphones,” *Int. J. Hum. Comput. Interact.*, vol. 29, no. 9, pp. 604–618, Sep. 2013.

[67] D. Sandlin, G. McGwin, and C. Owsley, “Association between vision impairment and driving exposure in older adults aged 70 years and over: a population-based examination.,” *Acta Ophthalmol.*, vol. 92, no. 3, pp. e207–12, May 2014.

[68] V. Tadić, G. L. Hundt, S. Keeley, and J. S. Rahi, “Seeing it my way: living with childhood onset visual disability.,” *Child. Care. Health Dev.*, Jun. 2014.

[69] J. H. Shandiz, A. Derakhshan, A. Daneshyar, A. Azimi, H. O. Moghaddam, A. A. Yekta, S. H. H. Yazdi, and H. Esmaily, “Effect of cataract type and severity on visual acuity and contrast sensitivity.,” *J. Ophthalmic Vis. Res.*, vol. 6, no. 1, pp. 26–31, Jan. 2011.

[70] D. B. Elliott and P. Situ, "Visual acuity versus letter contrast sensitivity in early cataract," *Vision Res.*, vol. 38, no. 13, pp. 2047–2052, Jun. 1998.

[71] M. S. Lasa, M. B. Datiles, M. J. Podgor, and B. V Magno, "Contrast and glare sensitivity. Association with the type and severity of the cataract.," *Ophthalmology*, vol. 99, no. 7, pp. 1045–9, Jul. 1992.

[72] C. Stephanidis and M. Antona, Eds., *Universal Access in Human-Computer Interaction. Applications and Services for Quality of Life*, vol. 8011. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013.

[73] F. Wilkinson, H. R. Wilson, and C. Habak, "Detection and recognition of radial frequency patterns.," *Vision Res.*, vol. 38, no. 22, pp. 3555–68, Nov. 1998.

[74] B. G. Jeffrey, Y. Z. Wang, and E. E. Birch, "Circular contour frequency in shape discrimination.," *Vision Res.*, vol. 42, no. 25, pp. 2773–9, Nov. 2002.

[75] C. T. M. de Castro, C. S. Kallie, and S. R. Salomão, "[Development and validation of the MNREAD reading acuity chart in Portuguese].," *Arq. Bras. Oftalmol.*, vol. 68, no. 6, pp. 777–83.

[76] A. Colenbrander, "Preservation of vision or prevention of blindness?," *Am. J. Ophthalmol.*, vol. 133, no. 2, pp. 263–265, Feb. 2002.

[77] August Colenbrander, "The COLENBRANDER Low Vision Measurement System." .

[78] P. Vision, "Visual Acuity." [Online]. Available: http://precision-vision.com/Introduction-to-Visual-Acuity-Measurement/a-visualacuity.html#.U6GgsfldVMK.

[79] E. Nielsen and J. Hjortdal, "Visual acuity and contrast sensitivity after posterior lamellar keratoplasty.," *Acta Ophthalmol.*, vol. 90, no. 8, pp. 756–60, Dec. 2012.

[80] S. a Haymes, K. F. Roberts, A. F. Cruess, M. T. Nicolela, R. P. LeBlanc, M. S. Ramsey, B. C. Chauhan, and P. H. Artes, "The letter contrast sensitivity test: clinical evaluation of a new design.," *Invest. Ophthalmol. Vis. Sci.*, vol. 47, no. 6, pp. 2739–45, Jun. 2006.

[81] P. S. Moke, a H. Turpin, R. W. Beck, J. M. Holmes, M. X. Repka, E. E. Birch, R. W. Hertle, R. T. Kraker, J. M. Miller, and C. a Johnson, "Computerized method of visual acuity testing: adaptation of the amblyopia treatment study visual acuity testing protocol.," *Am. J. Ophthalmol.*, vol. 132, no. 6, pp. 903–909, 2001.

[82] C. Introduction, S. Equipment, and C. R. Standards, "British Society of Audiology Recommended procedure Pure tone air and bone conduction threshold audiometry with and without masking and determination of uncomfortable loudness levels," no. March, pp. 1–27, 2004.

[83] Joe Walter Kutz, "Audiology Pure-Tone Testing." [Online]. Available: http://emedicine.medscape.com/article/1822962-overview.

[84] C. M. Rabelo and E. Schochat, "Time-compressed speech test in Brazilian Portuguese.," *Clinics (Sao Paulo).*, vol. 62, no. 3, pp. 261–72, Jun. 2007.

[85] A. M. Bishop, "Psychometrically Equivalent Bisyllabic WordLists for Word Recognition Testing in Spanish," no. August, 2009.

[86] T. Corina, S. De Angelo, M. C. Bevilacqua, A. Lima, and M. Moret, "Percepção da fala em deficientes auditivos pré-linguais usuários de implante coclear **** Speech perception in pre-lingual deaf users of cochlear implant," vol. 22, no. 3, pp. 275–279, 2010.

[87] P. Weisleder and W. R. Hodgson, "Evaluation of four Spanish word-recognition-ability lists.," *Ear Hear.*, vol. 10, no. 6, pp. 387–92, Dec. 1989.

[88] A. S.-L.-H. A. (ASHA), "Determining Threshold Level for Speech," American Speech-Language-Hearing Association, 1988.

[89] J. R. Schoepflin, "Back to Basics: Speech Audiometry." [Online]. Available: http://www.audiologyonline.com/articles/back-to-basics-speech-audiometry-6828.

[90] "The Audiogram."

[91] W. D. Center, "Understanding and Interpreting Lux Values." [Online]. Available: http://msdn.microsoft.com/en-us/library/windows/desktop/dd319008(v=vs.85).aspx.

[92] P. Schlyter, "Radiometry and photometry in astronomy." [Online]. Available: http://stjarnhimlen.se/comp/radfaq.html#10.

[93] J. Brooke, "SUS: A Retrospective," *J. Usability Stud.*, 2013.

[94] A. Bangor, P. Kortum, and J. Miller, "Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale," *J. Usability Stud.*, 2009.

[95] J. Sauro, "Measuring Usability With The System Usability Scale (SUS)," 2011. [Online]. Available: http://www.measuringusability.com/sus.php.

Appendix A

This Appendix contains information related to the testing procedures and the volunteers involved in the validation process mentioned in the User testing and Validation chapter.

7.1 Volunteers Information

User 0002

Idade	73
Problema de visão	Sim
Problema de audição	Não
Apoio à visão ou audição	Óculos
Usou smartphones	Sim (outros testes)
Costuma usar smartphones	Não
Tem um Smartphone	Não
Historico de queda	Não

Eye – Aided	Distance Visual Acuity:	Reading Visual Acuity:	Letter Visual Acuity:
Right – Yes	ND	0.4	0.93
Right – No	ND	ND	ND
Left – Yes	ND	1	0.8
Left – No	ND	ND	ND

User 0003

Idade	75
Problema de visão	Sim (Direito melhor que o esquero), problema no olho esquerdo
Problema de audição	Sim
Apoio à visão ou audição	Óculos
Usou smartphones	Sim
Costuma usar smartphones	Não
Tem um Smartphone	Não
Historico de queda	Não

Eye – Aided	Distance Visual Acuity:	Reading Visual Acuity:	Letter Visual Acuity:
Right – Yes	0.9	ND	0.625
Right – No	0.6	ND	0.25
Left – Yes	0.4	ND	0.2
Left – No	0.3	ND	0.125

User 0004

Idade	70
Problema de visão	Óculos
Problema de audição	Não
Apoio à visão ou audição	Óculos
Usou smartphones	Sim
Costuma usar smartphones	Não
Tem um Smartphone	Não
Historico de queda	Não

Eye – Aided	Distance Visual Acuity:	Reading Visual Acuity:	Letter Visual Acuity:
Right – Yes	0.7	1	0.4
Right – No	0.7	0.5	0.3
Left – Yes	0.7	0.63	0.3
Left – No	0.6	0.4	0.3

User 0005

Idade	66
Problema de visão	Sim (Direito melhor que Esquerdo)
Problema de audição	Não
Apoio à visão ou audição	Óculos
Usou smartphones	Sim
Costuma usar smartphones	Não
Tem um Smartphone	Não
Historico de queda	Sim (não muito frequentemente)

Eye - Aided	Distance Visual Acuity:	Reading Visual Acuity:	Letter Visual Acuity:
Right - Yes	0.9	ND	0.3
Right - No	0.8	ND	0.3
Left - Yes	0.5	ND	0.3
Left - No	0.3	ND	0.3

User 0006

Idade	71
Problema de visão	Sim (Esquerdo melhor que Direito)

Problema de audição	Não
Apoio à visão ou audição	Óculos
Usou smartphones	Sim
Costuma usar smartphones	Não
Tem um Smartphone	Não
Historico de queda	Não

Eye - Aided	Distance Visual Acuity:	Reading Visual Acuity:	Letter Visual Acuity:
Right - Yes	ND	1	0.5
Right - No	0.6	0.4	0.5
Left - Yes	ND	1	1
Left – No	1	0.4	0.5

User 0007

Idade	69
Problema de visão	Sim
Problema de audição	Sim (Esquerdo não houve nada. Direito usa aparelho)
Apoio à visão ou audição	Óculos e aparelho auditivo no ouvido direito
Usou smartphones	Sim
Costuma usar smartphones	Não
Tem um Smartphone	Não
Historico de queda	Não

Eye - Aided	Distance Visual Acuity:	Reading Visual Acuity:	Letter Visual Acuity:
Right - Yes	0.6	0.5	0.4
Right - No	0.5	0.63	0.3
Left - Yes	0.5	0.63	0.5
Left - No	0.2	0.4	0.4

User 0008

Idade	64
Problema de visão	Sim
Problema de audição	Aparelho (não trouxe)
Apoio à visão ou audição	Óculos, aparelho auditivo (mas não usa)
Usou smartphones	Sim
Costuma usar smartphones	Sim
Tem um Smartphone	Sim
Historico de queda	Sim

Eye - Aided	Distance Visual Acuity:	Reading Visual Acuity:	Letter Visual Acuity:
Right - Yes	0.5	ND	ND
Right - No	0.4	0.4	0.5
Left - Yes	0.7	ND	ND
Left - No	0.3	0.4	0.5

User 0009

Idade	71
Problema de visão	Sim (Direito melhor que esq)
Problema de audição	Não
Apoio à visão ou audição	Óculos
Usou smartphones	Sim - Não
Costuma usar smartphones	Não
Tem um Smartphone	Não
Historico de queda	Não (Já teve)

Eye - Aided	Distance Visual Acuity:	Reading Visual Acuity:	Letter Visual Acuity:
Right - Yes	ND	0.8	0.8
Right - No	0.8	0.4	0.3
Left - Yes	ND	0.32	0.3
Left - No	0.3	0.2	0.2

User 0010

Idade	70
Problema de visão	Sim (Direito melhor que Esquerdo)
Problema de audição	Não
Apoio à visão ou audição	Óculos
Usou smartphones	Sim
Costuma usar smartphones	Não
Tem um Smartphone	Não
Historico de queda	Não

Eye - Aided	Distance Visual Acuity:	Reading Visual Acuity:	Letter Visual Acuity:
Right - Yes	0.9	1	0.8
Right - No	ND	ND	ND
Left - Yes	1	1	0.8
Left - No	ND	ND	ND

User 0011

Idade	76
Problema de visão	Sim (Direito não vê)
Problema de audição	Não
Apoio à visão ou audição	Óculos
Usou smartphones	Sim
Costuma usar smartphones	Não
Tem um Smartphone	Não
Historico de queda	Não

Eye - Aided	Distance Visual Acuity:	Reading Visual Acuity:	Letter Visual Acuity:
Right - Yes	ND	ND	ND
Right - No	ND	ND	ND
Left - Yes	ND	0.8	0.625
Left - No	0.6	0.63	0.625

7.2 Volunteer Test Results

This section list the results obtained by each user in each test. For near and distance vision tests, if the user wore glasses the score will be followed by " – AIDED". Likewise, for user 0007 who took a PTA test with and without his hearing aid, the one in which he used it is also marked with "- AIDED".

User 0002

Reading Visual Acuity

Side:LEFT --- Score:0.05 - AIDED

Side:RIGHT --- Score:0.4 - AIDED

Contrast Sensitivity

Side:RIGHT --- Weber Contrast:1.5686276

Side:LEFT --- Weber Contrast:1.5686276

Shape Discrimination

Side:RIGHT --- Best Ram:0.011

Side:LEFT --- Best Ram:0.044

Speech Disc

Side:BOTH --- Score:0.95

Audiogram Press

Side:RIGHT

---> Frequency:250.0 - Volume:2/15

---> Frequency:500.0 - Volume:1/15

---> Frequency:1000.0 - Volume:1/15

---> Frequency:2000.0 - Volume:1/15

---> Frequency:4000.0 - Volume:3/15

---> Frequency:8000.0 - Volume:4/15

Side:LEFT

---> Frequency:250.0 - Volume:1/15

---> Frequency:500.0 - Volume:1/15

---> Frequency:1000.0 - Volume:1/15

---> Frequency:2000.0 - Volume:1/15

---> Frequency:4000.0 - Volume:1/15

---> Frequency:8000.0 - Volume:6/15

User 0003

Reading Visual Acuity

Side:LEFT --- Score:0.1

Side:LEFT --- Score:0.16 - AIDED

Side:RIGHT --- Score:0.25

Side:RIGHT --- Score:0.5 - AIDED

Contrast Sensitivity

Side:RIGHT --- Weber Contrast:1.1764706

Side:LEFT --- Weber Contrast:2.3529413

Shape Discrimination

Side:RIGHT --- Best Ram:0.02

Side:RIGHT --- Best Ram:0.008 - AIDED

Side:LEFT --- Best Ram:0.032

Side:LEFT --- Best Ram:0.02 - AIDED

Dist Visual Acuity

Side:LEFT --- Score:0.3

Side:LEFT --- Score:0.3 - AIDED

Side:RIGHT --- Score:0.3

Side:RIGHT --- Score:0.75 - AIDED

Speech Disc

Side:BOTH --- Score:0.95

Audiogram Press

Side:RIGHT

---> Frequency:250.0 - Volume:3/15

---> Frequency:500.0 - Volume:1/15

---> Frequency:1000.0 - Volume:1/15

---> Frequency:2000.0 - Volume:1/15

---> Frequency:4000.0 - Volume:2/15

---> Frequency:8000.0 - Volume:6/15

Side:LEFT

---> Frequency:250.0 - Volume:1/15

---> Frequency:500.0 - Volume:1/15

---> Frequency:1000.0 - Volume:1/15

---> Frequency:2000.0 - Volume:1/15

---> Frequency:4000.0 - Volume:3/15

---> Frequency:8000.0 - Volume:8/15

User 0004

Reading Visual Acuity

Side:RIGHT --- Score:0.25

Side:LEFT --- Score:0.125

Side:LEFT --- Score:0.4 - AIDED

Side:RIGHT --- Score:0.32 - AIDED

Contrast Sensitivity

Side:RIGHT --- Weber Contrast:2.745098

Side:LEFT --- Weber Contrast:5.098039

Shape Discrimination

Side:RIGHT --- Best Ram:0.026

Side:LEFT --- Best Ram:0.05

Dist Visual Acuity

Side:RIGHT --- Score:0.19230768

Side:LEFT --- Score:0.6

Side: LEFT --- Score:0.6 - AIDED

Side:RIGHT --- Score:0.3 - AIDED

Speech Disc

Side:BOTH --- Score:1.0

Audiogram Press

Side:RIGHT

---> Frequency:250.0 - Volume:2/15

---> Frequency:500.0 - Volume:1/15

---> Frequency:1000.0 - Volume:2/15

---> Frequency:2000.0 - Volume:1/15

---> Frequency:4000.0 - Volume:1/15

---> Frequency:8000.0 - Volume:2/15

Side:LEFT

---> Frequency:250.0 - Volume:2/15

---> Frequency:500.0 - Volume:1/15

---> Frequency:1000.0 - Volume:1/15

---> Frequency:2000.0 - Volume:1/15

---> Frequency:4000.0 - Volume:1/15

---> Frequency:8000.0 - Volume:1/15

User 0005

Reading Visual Acuity

Side:LEFT --- Score:0.2

Side:RIGHT --- Score:0.2

Side:LEFT --- Score:0.4

Side:RIGHT --- Score:0.5

Contrast Sensitivity

Side:RIGHT --- Weber Contrast:1.9607844

Side:LEFT --- Weber Contrast:1.5686276

Shape Discrimination

Side:LEFT --- Best Ram:0.038

Side:RIGHT --- Best Ram:0.032

Side:LEFT --- Best Ram:0.026

Side:RIGHT --- Best Ram:0.032

Dist Visual Acuity

Side:RIGHT --- Score:0.3

Side:LEFT --- Score:0.3

Side:LEFT --- Score:0.6 - AIDED

Side:RIGHT --- Score:0.75 - AIDED

Speech Disc

Side:BOTH --- Score:0.95

Audiogram Press

Side:RIGHT

---> Frequency:250.0 - Volume:2/15

---> Frequency:500.0 - Volume:1/15

---> Frequency:1000.0 - Volume:1/15

---> Frequency:2000.0 - Volume:1/15

---> Frequency:4000.0 - Volume:1/15

---> Frequency:8000.0 - Volume:1/15

Side:LEFT

---> Frequency:250.0 - Volume:1/15

---> Frequency:500.0 - Volume:1/15

---> Frequency:1000.0 - Volume:1/15

---> Frequency:2000.0 - Volume:1/15

---> Frequency:4000.0 - Volume:1/15

---> Frequency:8000.0 - Volume:1/15

User 0006

Reading Visual Acuity

Side:RIGHT --- Score:0.25

Side:LEFT --- Score:0.25

Side:RIGHT --- Score:0.25 - AIDED

Side:LEFT --- Score:0.5 - AIDED

Contrast Sensitivity

Side:LEFT --- Weber Contrast:1.5686276

Side:RIGHT --- Weber Contrast:1.9607844

Shape Discrimination

Dist Visual Acuity

Side:RIGHT --- Score:0.6

Side:LEFT --- Score:1.2

Speech Disc

Side:BOTH --- Score:0.85

Audiogram Press

Side:RIGHT

---> Frequency:250.0 - Volume:2/15

---> Frequency:500.0 - Volume:1/15

---> Frequency:1000.0 - Volume:2/15

---> Frequency:2000.0 - Volume:1/15

---> Frequency:4000.0 - Volume:5/15

---> Frequency:8000.0 - Volume:3/15

Side:LEFT

---> Frequency:250.0 - Volume:1/15

---> Frequency:500.0 - Volume:1/15

---> Frequency:1000.0 - Volume:1/15

---> Frequency:2000.0 - Volume:2/15

---> Frequency:4000.0 - Volume:3/15

---> Frequency:8000.0 - Volume:3/15

User 0007

Reading Visual Acuity

Side:RIGHT --- Score:0.32

Side:LEFT --- Score:0.25

Side:LEFT --- Score:0.32 - AIDED

Side:RIGHT --- Score:0.25 - AIDED

Contrast Sensitivity

Side:RIGHT --- Weber Contrast:1.1764706

Side:LEFT --- Weber Contrast:1.5686276

Shape Discrimination

Side:LEFT --- Best Ram:0.011

Side:RIGHT --- Best Ram:0.014

Dist Visual Acuity

Side:RIGHT --- Score:0.6

Side:LEFT --- Score:0.3

Side:LEFT --- Score:0.6 - AIDED

Side:RIGHT --- Score:0.3 - AIDED

Speech Disc

Side:BOTH --- Score:0.95

Audiogram Press

Side:RIGHT - AIDED

---> Frequency:250.0 - Volume:1/15

---> Frequency:500.0 - Volume:1/15

---> Frequency:1000.0 - Volume:1/15

---> Frequency:2000.0 - Volume:3/15

---> Frequency:4000.0 - Volume:2/15

---> Frequency:8000.0 - Volume:6/15

Side:LEFT

---> Frequency:250.0 - Volume:7/15

---> Frequency:500.0 - Volume:7/15

---> Frequency:1000.0 - Volume:6/15

---> Frequency:2000.0 - Volume:8/15

---> Frequency:4000.0 - Volume:Não Ouviu!

---> Frequency:8000.0 - Volume:7/15

Side:RIGHT

---> Frequency:250.0 - Volume:1/15

---> Frequency:500.0 - Volume:1/15

---> Frequency:1000.0 - Volume:1/15

---> Frequency:2000.0 - Volume:7/15

---> Frequency:4000.0 - Volume:8/15

---> Frequency:8000.0 - Volume:8/15

User 0008

Reading Visual Acuity

Side:RIGHT --- Score:0.25

Side:LEFT --- Score:0.32

Contrast Sensitivity

Side:RIGHT --- Weber Contrast:1.9607844

Side:LEFT --- Weber Contrast:1.9607844

Shape Discrimination

Side:RIGHT --- Best Ram:0.014

Side:LEFT --- Best Ram:0.026

Dist Visual Acuity

Side:RIGHT --- Score:0.3

Side:LEFT --- Score:0.24

Side:RIGHT --- Score:0.3 - AIDED

Side:LEFT --- Score:0.6 - AIDED

Speech Disc

Side:BOTH --- Score:0.8

Audiogram Press

Side:RIGHT

---> Frequency:250.0 - Volume:5/15

---> Frequency:500.0 - Volume:4/15

---> Frequency:1000.0 - Volume:3/15

---> Frequency:2000.0 - Volume:2/15

---> Frequency:4000.0 - Volume:7/15

---> Frequency:8000.0 - Volume:Não Ouviu!

Side:LEFT

---> Frequency:250.0 - Volume:1/15

---> Frequency:500.0 - Volume:1/15

---> Frequency:1000.0 - Volume:1/15

---> Frequency:2000.0 - Volume:5/15

---> Frequency:4000.0 - Volume:5/15

---> Frequency:8000.0 - Volume:Não Ouviu!

User 0009

Reading Visual Acuity

Side:LEFT --- Score:0.1

Side:RIGHT --- Score:0.2

Side:LEFT --- Score:0.125 - AIDED

Side:RIGHT --- Score:0.5 - AIDED

Contrast Sensitivity

Side:LEFT --- Weber Contrast:3.9215689

Side:RIGHT --- Weber Contrast:1.9607844

Shape Discrimination

Side:LEFT --- Best Ram:0.02

Side:RIGHT --- Best Ram:0.02

Dist Visual Acuity

Side:LEFT --- Score:0.3

Side:RIGHT --- Score:0.75

Speech Disc

Side:BOTH --- Score:0.95

Audiogram Press

Side:RIGHT

---> Frequency:250.0 - Volume:1/15

---> Frequency:500.0 - Volume:1/15

---> Frequency:1000.0 - Volume:1/15

---> Frequency:2000.0 - Volume:1/15

---> Frequency:4000.0 - Volume:1/15

---> Frequency:8000.0 - Volume:1/15

Side:LEFT

---> Frequency:250.0 - Volume:1/15

---> Frequency:500.0 - Volume:1/15

---> Frequency:1000.0 - Volume:1/15

---> Frequency:2000.0 - Volume:1/15

---> Frequency:4000.0 - Volume:1/15

---> Frequency:8000.0 - Volume:2/15

User 0010

Reading Visual Acuity

Side:LEFT --- Score:0.5

Side:RIGHT --- Score:0.5

Contrast Sensitivity

Side:BOTH --- Weber Contrast:1.5686276

Shape Discrimination

Side:RIGHT --- Best Ram:0.026

Side:LEFT --- Best Ram:0.02

Dist Visual Acuity

Side:RIGHT --- Score:0.75

Side:LEFT --- Score:0.9375

Speech Disc

Side:BOTH --- Score:1.0

Audiogram Press

Side:RIGHT

---> Frequency:0.0 - Volume:Não Ouviu!

---> Frequency:250.0 - Volume:1/15

---> Frequency:500.0 - Volume:1/15

---> Frequency:1000.0 - Volume:2/15

---> Frequency:2000.0 - Volume:1/15

---> Frequency:4000.0 - Volume:1/15

---> Frequency:8000.0 - Volume:3/15

Side:LEFT

---> Frequency:0.0 - Volume:Não Ouviu!

---> Frequency:250.0 - Volume:1/15

---> Frequency:500.0 - Volume:1/15

---> Frequency:1000.0 - Volume:1/15

---> Frequency:2000.0 - Volume:1/15

---> Frequency:4000.0 - Volume:1/15

---> Frequency:8000.0 - Volume:4/15

User 0011

Reading Visual Acuity

Side:LEFT --- Score:0.32

Side:LEFT --- Score:0.32 - AIDED

Contrast Sensitivity

Side:LEFT --- Weber Contrast:1.9607844

Shape Discrimination

Side:LEFT --- Best Ram:0.014

Side:LEFT --- Best Ram:0.02

Dist Visual Acuity

Side:LEFT --- Score:0.6

Speech Disc

Side:BOTH --- Score:0.85

Audiogram Press

Side:RIGHT

---> Frequency:0.0 - Volume:Não Ouviu!

---> Frequency:250.0 - Volume:2/15

---> Frequency:500.0 - Volume:1/15

---> Frequency:1000.0 - Volume:1/15

---> Frequency:2000.0 - Volume:3/15

---> Frequency:4000.0 - Volume:6/15

---> Frequency:8000.0 - Volume:Não Ouviu!

Side:LEFT

---> Frequency:0.0 - Volume:Não Ouviu!

---> Frequency:250.0 - Volume:2/15

---> Frequency:500.0 - Volume:1/15

---> Frequency:1000.0 - Volume:1/15

---> Frequency:2000.0 - Volume:3/15

---> Frequency:4000.0 - Volume:4/15

---> Frequency:8000.0 - Volume:Não Ouviu!

7.3 Task Specific Guidelines

Before starting a game related task, the user would be given some guidelines on how to interact with it. Those guidelines are presented below.

Visão ao Perto (Task 2):

O teste de Visão ao Perto pretende testar a sua visão ao perto, assim sendo peço-lhe que não se aproxime nem se afaste do dispositivo, no entanto terá que rodar o dispositivo de modo a conseguir ler o texto. Ser-lhe à apresentada uma pequena frase, nesta frase algumas letras estão escondidas atrás de um quadrado vermelho. Pela ordem de leitura normal, da esquerda para a direita e de cima para baixo, deve selecionar a letra em falta clicando no botão refente a essa letra. Se não tiver a certeza da letra em falta ou houver mais que uma letra correcta, escolha a que lhe parecer mais correcta. Quando acerta numa letra, esta torna-se visivel e deve repetir o processo para a próxima letra. Quando descobrir todas as letras, ser-lhe á dado uma nova frase com um tamanho mais pequeno. Isto repetir-se-á até que não consiga mais ler o texto e falhe demasiadas tentativas, aí terá completado a sua tarefa e ser-lhe-á apresentado o seu resultado.

Relembro que deve manter o dispositivo sempre à mesma distância da sua cara.

Contraste (Task 3):

O teste de Contraste pretende testar a sensibilidade à mudança de cores., assim sendo peço-lhe que não se aproxime nem se afaste do dispositivo. Um C ser-lhe-á apresentado numa certa rotação. Deve carregar na maçã para onde a abertura do C se encontra virada, deste modo o C poderá comer aquela maçã. À medida que responde corretamente o C vai ficando mais claro até que se torna praticamente invisivel. Quando não tiver a certeza da rotação certa, escolha a que lhe parecer mais correta. Se não conseguir mesmo ver o C, escolha uma orientação à sorte.

Isto repetir-se-á até que a sua sensibilidade seja definida, aí terá completado a sua tarefa e ser-lhe-á apresentado o seu resultado.

Relembro que deve manter o dispositivo sempre à mesma distância da sua cara.

Identificação de Palavras (Task 5):

O teste de Identificação de Palavras pretende testar a sua capacidade de identificar palavras que são semelhantes à outras. Com os auscultadores nos ouvidos vai ouvir uma palavra numa voz de uma mulher com um sotaque brasileiro. Vão lhe ser apresentadas quatro (4) respostas possíveis, deve escolher a palavra que a mulher disse. Caso não tenha a certeza deve escolher a que lhe parecer mais correta.

Isto repetir-se à durante 20 tentativas, no final ser-lhe-á apresentado um resultado e a sua tarefa estará terminada.

Detecção de Sons (Task 6):

O teste de Deteção de Sons pretende testar a sua capacidade de detectar sons a diferentes frequências. Com os auscultadores nos ouvidos vai -lhe ser apresentado um animal dentro de um quadrado. A certas alturas vai ouvir um som, quando isto acontecer deve pressionar e manter pressionado o animal no interior deste quadrado, até que deixe de ouvir o som. O animal e o quadrado estarão sempre presentes, quer haja som quer não mas só deve pressionar quando ouvir o som. Quando deixar de ouvir o som pode tirar o dedo do ecrã.

Quando o teste acabar, ser-lhe-á apresentado um resultado e a sua tarefa estará terminada.

Comparação de Formas (Task 9):

O teste de Comparação de Formas pretende testar a sua capacidade de identificar formas iguais, assim sendo peço-lhe que não se aproxime nem se afaste do dispositivo, no entanto poderá rodar o dispositivo se assim o desejar . Ser-lhe-ão apresentadas seis (6) formas. Estas formas eram inicialmente circulos que foram deformados de maneira a ficarem cada vez mais diferentes. Neste exemplo (Figure 51 is shown), a forma mais deformada é a da esquerda e a deformidade vai sendo reduzida da esquerda para a direita. Cada forma tem outra igual a sua, desta forma tem que fazer três (3) pares de duas formas. É importante começar por fazer os pares mais deformados e só depois passar aos menos deformados, deve fazê-lo por ordem do mais deformado ao menos deformado. Se não tiver a certeza de qual a próxima forma a escolher, escolha a que achar mais correta, caso escolha a errada será avisado deste facto.

Isto repetir-se-á até a sua sensibilidade seja definida, aí terá completado a sua tarefa e ser-lhe-á apresentado o seu resultado.

Relembro que deve manter o dispositivo sempre à mesma distância da sua cara.

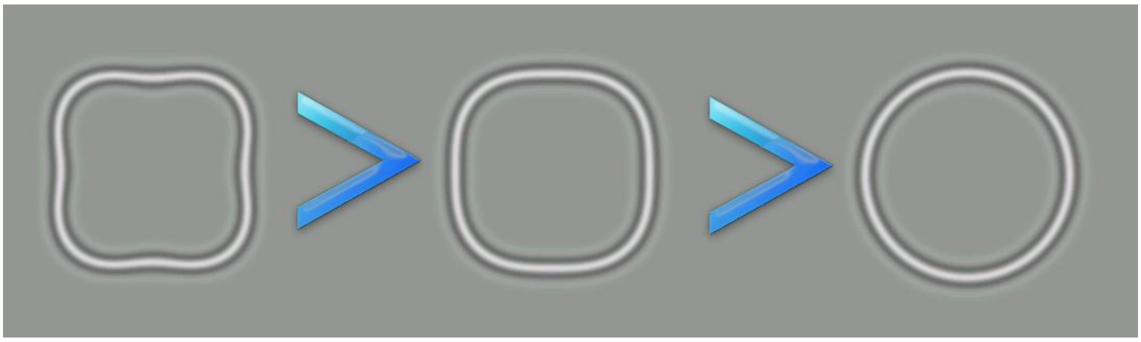

Figure 51. Shape Discrimination Explanation

Visão ao Longe (Task 10):

O teste de Visão ao Longe pretende testar a sua visão ao longe. Deve colocar-se a um metro e meio (1.5m) do espelho, no local indicado. Deve ter o dispositivo junto ao seu corpo e virar o ecrã do dispositivo de frente para o espelho. Deve olhar apenas para o reflexo do dispositivo no espelho, nunca pode olhar diretamente para o dispositivo. No dispositivo aparecerá um E com a sua abertura para cima, baixo, esquerda ou direita. O seu objectivo é rodar o dispositivo de modo a que a abertura do E se encontre a apontar para o tecto. Quando achar que tem o dispositivo na posição certa, deve pressionar o ecrã até que este trema e mude de posição, indicando um novo teste que deve completar da mesma forma. Se não tiver a certeza quanto à resposta certa, escolha a que lhe parecer mais correta.

Deve repetir este procedimento até que lhe seja apresentado um resultado em vez de um teste novo.

Relembro que deve manter o dispositivo sempre à mesma distância do espelho e que nunca deve olhar diretamente para o espelho até o jogo estar completado.

7.4 SUS Questionnaire

The users were instructed to use the questionnaire presented below and select the option they thought best represented their experience with the system being evaluated.

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	1	2	3	4	5	Concordo

fortemente						fortemente

Qual foi a maior dificuldade que encontrou em usar este sistema?

____

____

____

____

____

Tem alguma recomendação ou funcionalidade que gostaria de ver no sistema?

____

____

____

____

____