An Android application for helping in the identification of Children with Reading Difficulties

Rita Francese
University of Salerno
Fisciano (SA), Italy
francese@unisa.it

Clara Monaco
University of Salerno
Fisciano (SA), Italy
c.monaco16@studenti.unisa.it

Claudia Nicoletti
Fondazione Salernum Anffas Onlus
Salerno, Italy
psicologac@anffas.sa.it

ABSTRACT

It is very important that children with reading disorders be detected in their early age because this problem may have a hard impact on the learning performance and cause study drop-out, lower self-esteem, and, in extreme case, unemployment. In this paper we present an application aiming at providing help in the identification of 7/8-year-old Italian children attending the second class of the primary school having reading disorders. The app is a reading exercise articulated in levels and is based on sentences created on purpose. Specific metrics for detecting reading problems are also proposed and the threshold values are experimentally set. The sample is composed of 46 second class children of a primary school. A preliminary validation of the application has been conducted on a little sample of children with learning difficulties. First results are encouraging and suggest that the application may be tested on a wider scale.

CCS CONCEPTS

• **General and reference** → **Experimentation**; • **Software and its engineering** → Interactive games;

KEYWORDS

Dyslexia, Disability, Mobile Application, Interactive Game

ACM Reference Format:
Rita Francese, Clara Monaco, and Claudia Nicoletti. 2018. An Android application for helping in the identification of Children with Reading Difficulties. In *International Conference on Smart Objects and Technologies for Social Good (Goodtechs '18), November 28–30, 2018, Bologna, Italy*. ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3284869.3284915

1 INTRODUCTION

Reading difficulties can have different origins. One of them is Dyslexia, a neurodevelopmental disorder belonging to the Specific Learning Disabilities (SLD) that impedes to learn or use specific academic skills[1]. Early signs of reading difficulties may appear in the preschool years (e.g., difficulty in learning names of letters or reading aloud), but they can only be diagnosed reliably after starting formal education. Typically, people with Dyslexia have an IQ in the norm, have difficulties in reading aloud, in spelling word, in writing, often also difficulty with the system of numbers and calculation. Sometimes they have difficulties in reading comprehension of the text, in oral language and in vocabulary growth and background knowledge. The learning difficulties are 'unexpected' because other aspects of development seem to be fine. Indeed, generally, Dyslexia is not due to poor instruction or sensorial/neurological disease.

Dyslexia is frequent. In Italy the prevalence of ascertained dyslexic children varies between 3 and 5%, while in the English-speaking countries the percentage rises to 17%. It is very important for a person's life that these problems be detected in the early age because they may have a hard impact on the learning performance and cause study drop-out, lower self-esteem, and other serious consequences as unemployment.

In Italy there is a recent law (n. 170 of 2010) which recognizes the existence of SLD. The related implementation decree and guidelines (Decr. MIUR n.5669) state that each school, at the beginning of the school year, must detect and sign, together with the parents of diagnosed dyslexic children, the Personalized Learning Plan, obliging the School to implement all the necessary dispensative and compensatory methods and instruments for supporting the child during his/her learning activity. To get this support there is the need to have an early detection of this problem. It may be useful to empower the children parents providing them a way to perform a first check [23].

The use of serious games and the application of game elements may increase the intrinsic motivation: players refer to play for the pure enjoyment of game play [7, 18]. Gamification may successfully address this intrinsic motivation by using game mechanics to attract learners to the pure enjoyment of the activities. In addition, Generation Z[2] make a very large use of mobile applications, especially for gaming purpose and most dyslexic people prefer to use smartphones for learning [15]. This attitude may be used for investigating the reading ability of children in a nice way while their motivation is enforced by the use of the mobile device [19]. There are also recent computerized applications adopted in SLD rehabilitation (e.g., RiDiNet[3] or Tachidino[4]).

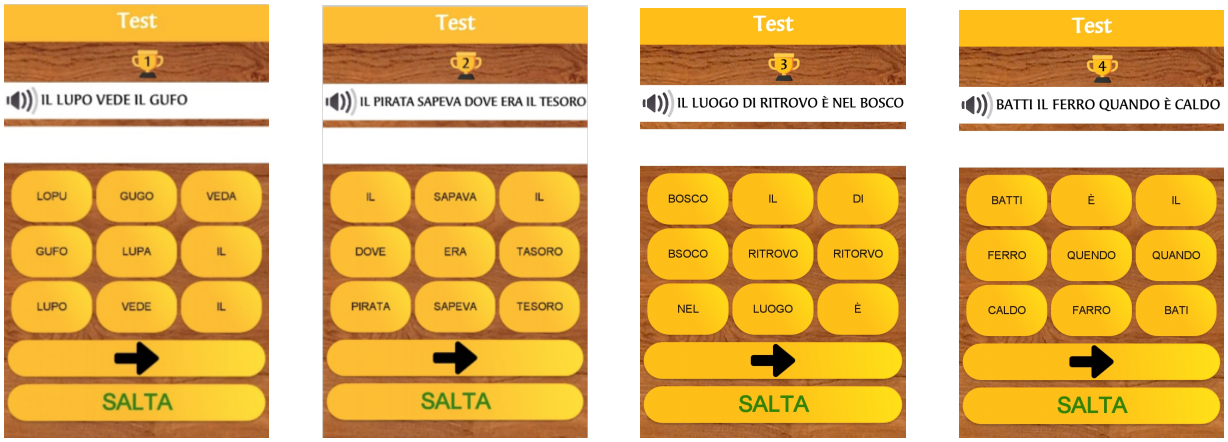
We here try to address this need by proposing a mobile application aiming at providing help in the identification of 7/8-year-old Italian children attending the second class of the primary school having reading disturbs. By providing a service that may benefit a

[1]International Dyslexia Association. Definition of Dyslexia: http://interdys.org/DyslexiaDefinition.htm

GOODTECHS, OCTOBER 2018, Bologna, Italy

ACM ISBN 978-1-4503-6581-9/18/11…$15.00
https://doi.org/10.1145/3284869.3284915

[2]The demographic cohort after the Millennials
[3]https://info.ridinet.it/ridinet-cosa/
[4]https://www.tachidino.com



(a) (b) (c) (d)

Figure 1: The four levels of the app.

large number of children, this app alerts that there may be some problems and the child needs an appropriate medical screening. The app application consists in reading 21 sentences created on purpose and based on the developmental model of acquiring literacy [8]. The app we propose is articulated in levels, in such a way as to avoid the frustration of having to answer too difficult questions. Specific metrics based on Information Retrival techniques (i..e., Precision, Recall and F-measure) have been proposed for measuring the reading performances. The thresholds needed to alert parents and teachers have been experimentally computed by proposing the app to a sample of 46 second class children of a primary school. Then, a preliminary validation has been conducted on a sample of children with learning problems.

The paper has been organized as follows: Section 2 discusses related work; Section 3 presents the app we propose; Section 4 describes the empirical setting of the alert thresholds and a preliminary evaluation involving real problematic cases, while Section 5 concludes the paper and discusses future work.

2 RELATED WORK

Many commercial tools and research work have been devoted to children's rehabilitation. When visiting Google play and using "Dislessia" and "Dyslexia" as search key we found more than 200 apps, but few of them are for Italian children. As an example, Double Quiz[5] is a game to learn to use double consonants in Italian. Many of them have been designed to conduct a rehabilitation program with children who have specific orthography disorders, such as Dislessia Evolutiva Pro [6]. Others are with screening purpose: for instance, [5] proposes only a questionnaire for a minute short story written in lowercase letters [21]. Also the App Store by Apple offers a lot of apps. A list of the most helpful may be found at *http://dyslexiahelp.umich.edu/tools/apps*. They are mainly devoted to therapy or provide help during reading activities.

Many approaches for assisting dyslexic people have been also proposed in literature, a systematic analysis may be found in [10]. As an example, Gaggi *et Al.* were focused in the early detection of Dyslexia in preschoolers and proposed and evaluated various serious games designed for predicting the risk of this disturb as soon as possible [9]. They investigated abilities related to Dyslexia, such as eye-hand coordination, visual search ability, rapid identification of visual and acoustic inputs and the capability to associate visual and auditory stimuli. [6] presents a tangible user interface proposing tactile letters. In [20] the authors discussed the rationale and protocol for the design and development of a digital health solution aimed at improving the early detection, monitoring and management of Dyslexia in young children (4-8 years). The project is at a preliminary phase. No description of the game modality is provided.

Many tools support reading activities, such as a Dyslexia-friendly reader app, which has been presented and evaluated in [13]. Madrigale [22] and Dyseggxia [1] employed pedagogical agents to support children in reading activities.

3 THE READING TEST

The app we propose aims at providing a first alert for children having reading problems. The test has no diagnosis purpose, because several condition may have a poor test results as a consequence, including a child's particular mood or intellectual developmental disability (IDD).

The test is composed of 21 sentences. Each sentence has an entire screen at disposal, as depicted in Fig. 1. The child has to reproduce the sentence by selecting the words which appear on the buttons. Each button shows a word of the sentence or a *distractor*, a word which is similar to the right one, but differs for a little particular and may create confusion to the reader. Disctractors differs from the right words by some letters or some syllabes. As an example in Fig. 1(a) the sentence is "IL LUPO VEDE IL GUFO", where the word "GUFO" may be exchanged with the word "GUGO", possible error of a dyslexic person. The test is organized in four levels, according to

[5]https://play.google.com/store/apps/details?id=it.pepesenise.doppiequiz
[6]https://play.google.com/store/apps/details?id=it.erickson.dislessiaevolutiva2

an encreasing length and complexity of the orthographic structure of the words:

- two-syllables (e.g., "Mara vede la luna");
- three-syllables (e.g., "il pirata lega le reti");
- complex two and three syllables (e.g.,"in vacanza porto il costume")
- many complex words with orthographic groups (e.g., "chi di spada ferisce di spada perisce").

The sentences are randomly selected. We collected 21 of them for each level, thus the pool of sentences is composed of 21*4 sentences, while the child has to answer 21 of them. These sentences and the associated distractors are stored in the application files. Together with a sentence the app randomly selects also the distractors associated with it. After four correct sentences of a given level the child moves to the next level. If he makes two errors he goes back to the previous level. The test ends when the child has examined all the selected 21 sentences.

At the end of the test results are provided in terms of an indication on the need of a neuropsychological diagnosis. All the mistakes are traced and can be detailedly examined.

The first version of the prototype has been validated by Dr. Nicoletti, one of the authors, expert in children with learning disorders. In particular, she noted that the interface adopted font with serif, which should be avoided to improve readability. Also the font on the buttons was too small and had a poor "readability". These suggestions have been implemented in the prototype version adopted during the rest of the paper.

4 EXPERIMENTAL ALERT THRESHOLD SETTINGS

When proposing a test aiming at alerting parents and teachers of specific learning problems of their children there is the need of accurately setting the variables on which performing the evaluation and the corresponding alert thresholds. Performance may change significantly with the age of the subject and in relation to the cognitive and neuropsycological profile of the child. It is possible to hypothesize that at different age different tests may be more reliable in detecting the disorder [4].

Concerning the minimum age in which the diagnosis can be made, it should theoretically coincide with the completion of the 2nd year of primary school (2nd grade), since this age coincides with the completion of the formal education cycle of the written code [4]. According to [4], there is the need of investigating possible problems when it occurs one (or both) of the following condition:

- reading speed rs: there exists a significant distance from the expected average values for the class attended by the child, conventionally set at $rs < \mu - 2 * \sigma$,
- reading accuracy: result under the 5th percentile.

The test we propose has the particularity to be organized in levels. So, we have to verify if the introductions of levels has to be considered as a relevant factor and try to set thresholds suitable for an online test. The design and planning of the experimental thresholds' setting has been structured according to the guidelines of Juristo and Moreno [11], Kitchenham *et al.* [12], and Wohlin *et al.* [24], suitable for empirical investigations. By Applying the Goal Question Metric (GQM) template [2] our goal can be defined as:

- *Analyze* the proposed reading test
- *for the purpose of* setting the alert thresholds
- *with respect to* the reading ability and the effort
- *from the viewpoint of* 7-8-year-old children attending the second class of a primary school
- *in the context of* their classroom.

4.1 Context Selection

The context of the experiment is the correct reading of sentences and words. The study was executed at a primary school of Salerno, Italy. The participants were 46 children of the second primary class. They were from two different classes, with a comparable background.

Before starting the test, the students were asked to fill in on the mobile device a pre-questionnaire aiming at collecting their demographic information, such as their gender and age.

Participants were informed that their performance in that experiment would not have any effect on their grades; moreover, their parents provided written consensus to the children's participation and were also informed that data collected during the experiment would only be exploited for research purposes, confidentially treated, and shared in anonymous and aggregated form.

4.2 Variable and threshold definition

Following the guidelines specified in [4], we have to signal children with problems in reading speed and/or accuracy. Thus, to analyze the participants' performances, we select the following variables adopted in literature for measuring Comprehension (see for example [17]) and we introduce the last for considering the effect of the reached level.

4.2.1 Effort. To assess the *Effort* variable we used the test completion time, which indicates the number of minutes to finish the test and then it assumes only positive integer values. Low values for task effort mean that participants spent less time (or less effort) to complete a reading task. The threshold is computed as: $\mu(Effort) + 2 * \sigma(Effort)$

4.2.2 Effectiveness. Effectiveness is evaluated in terms of the answers to the test provided by the participants. The test offers the possibility of skipping a question. So, for each question a child can give the right answer, the wrong answer or no answer at all. Effectiveness is computed in terms of *F-Measure* [16], an Information Retrieval metric. Given:

- n, the number of questions proposed during the test,
- m_s, the number of answers provided by the participant s,
- k_s, the number of correct answers participant s provides,

we compute the following measures: $precision_s = \frac{k_s}{m_s}$ and $recall_s = \frac{k_s}{n}$. Precision and recall measure two different concerns, namely correctness and completeness of the answers, respectively. To balance them, we adopted F-Measure, their harmonic mean (i.e., $\frac{2 \cdot precision_s \cdot recall_s}{precision_s + recall_s}$). This mean has been used to measure the *Effectiveness*. The measures above assume values in the range $[0, 1]$. For this measure, a

Figure 2: Two children during the treshold setting experiment.

value equal to 1 means that a participant correctly and completely answered all the test questions.

The threshold is computed considering the scores under the first percentile of Effectiveness. Similarly for the following two variables.

4.2.3 Efficiency. We also considered the efficiency with which the test was accomplished. It is a measure computed as the ratio between Effectiveness and test completion time and estimates the efficiency of a participant during a reading task. In particular, this measure estimates the ability of a participant to effectively answer the questions without wasting time:

$$Efficiency = \frac{Effectiveness}{Effort}$$

4.2.4 Efficiency by level. We decided to add a further measure: *Efficiency_by_level=Efficiency*Level*. This measure takes into account the maximum level reached by a participants (Level) during the test. As an example, this should prize a participant who has correctly answered 10 questions having reached the level 3 w.r.t. a participant answering the same number of question while remaining always at the first level.

4.3 Experiment planning and material

Before the experiment, each class received a brief description (5 minutes) of the app and of the experiment modalities by one of the authors. Participants played the app two at time, under the supervision of one of the authors (see Fig. 2). At the end of the test, the application annotated the ending time and all the answers. Then, the participants filled in the Post-Experiment Questionnaire, reported in Fig. 4. The questionnaire was included at the end of the test. It was composed of 8 questions expecting closed answers scored using the 5-point Likert scale [14] specified in Figure 3. We decided to use emoticons because their meaning is well known to children and they make the questionnaire more appealing.

The purpose of the Post-Experiment Questionnaire was to assess whether participants understood the app rules, the perceived test

Figure 3: The emoticons'Likert scale.

Table 1: Descriptive statistics for metrics.

Variable	Median	Mean	StDev	Min	Max	threshold
Effectiveness	0.649	0.617	0.153	0.214	0.905	0.52
Effort	9	9.565	3.160	4	20	15.89
Efficiency	0.065	0.071	0.028	0.024	0.146	0.046
Efficiency_by_Level	0.181	0.199	0.131	0.024	0.463	0.066

difficulty and their opinion on the app interface and usability [3]. Only few questions were proposed to avoid overloading.

4.4 Threshold setting

In the following we describe how we experimentally set the thresholds' values for the considered variables.

Descriptive statistics on Effort, Effectiveness, Efficiency and Efficiency_by_level are shown in Table 1. To further examine the threshold appropriateness we applied the k-means clustering algorithm to our dataset, with k=5, for all the variables except Effort. Results oscillated by varying the starting point between 9-10 elements in the fifth group. Thus, being $46/5 = 9.2$ we set the thresholds considering the last group of 10 elements. Considering Effort, only two children took more than 15.885 seconds: children 31 and 34. Child 34 has reached the best Effectiveness result in the test with the maximum Effort. The children with declared learning problems are numbered 24 (autistic) and 36 (dyslexic). They are both detected by all the measures except Effort. Children 31 and 34 are also detected by Efficiency. Efficiency by Level signals all the participants stopping at the first level.

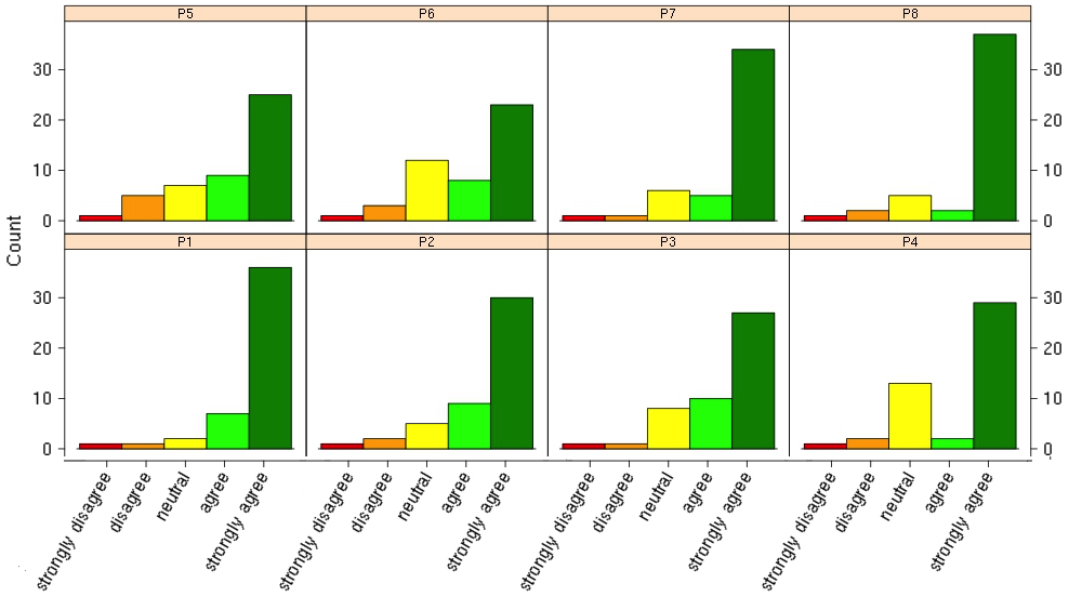
.

4.5 Children Perceptions

After playing children answered the eight questions reported in Fig. 4. The results collected by the app are shown in Fig. 4 in terms of an histogram for each question. In all the histogram there is a single strongly disagreeing opinion. It has been provided by a borderline autistic child who lamented not to understand the app, but that was not signaled by the metrics. Concerning the questions related to the test, its rules were perceived as clear by 43 participants (**P1**), it was considered attractive by 40 participants (**P2**) and the test duration was appropriate for 35 participants (**P5**). Concerning their carefulness (**P3**), 38 participants declared to carefully providing the answers, 5 were neutral. The reading questions were judged easy by 32 participants (**P4**), while 32 participants thought they have answered correctly (**P6**). Concerning the app usability, 40 children perceived the app as easy to use (**P7**) and the the buttons to be pressed were clear (**P8**).

4.6 Further validation

After the first experimentation in the primary school we began to collect other data for verifying the test validity and which variables



ID	Question	ID	Question
P1	The test rules were clear.	P5	The test duration was right.
P2	The app has attracted me.	P6	I think to have provided the right answers.
P3	I was careful when providing the answers.	P7	The app was easy to use.
P4	The questions were easy.	P8	It was clear which button I had to press.

Figure 4: The Post-Experiment questionnaire results.

better detect students with reading problems. In particular, after having had the parents' consensus, Dr. Nicoletti proposed the app to her patients, who suffer all of learning disturbs, but with different pathologies and age. At the present, nine children used the app. The collected metrics are reported in Table 2, together with the participant's age and a description of his/her disturb; the measures in bold highlight that the participant is under the corresponding threshold. id8 and id9 in Table 2 were detected by all the variables except Effort. Also id3 was detected, who was ten years old, but frequenting the second primary class. Effectiveness also detected id6, who was nine years old. This child was very fast but with bad performances.

4.7 Discussion

The preliminary results shown in Table 2 are encouraging and seems to confirm the thresholds' validity in case of seven-year-old children. All the considered metrics identify the already signaled children attending the second primary class. A large scale experiment should better investigate which of them are more suitable or if they are all relevant.

The results of the Post-Experiment questionnaire made us make a reflection on the number of questions. The number we selected, 21 questions, and the individuated levels seems to be appropriate for detecting reading difficulties, but 5 participants expressed a negative judgment on the test duration. To investigate if it is the case of reducing the test number of questions or to consider a different number depending on the participants ability we examined the correlation coefficient between the answers provided to question P5 and the corresponding Effectiveness. The Pearson correlation coefficient between the opinion on question P5 and Effectiveness is -0.17, very small. This means that there is a negligible relationship (inverse) between the performances and the perception of the test duration. So we decided to leave 21 questions in the test because of the encouraging preliminary results.

5 CONCLUSIONS

In this paper we presented a mobile application aiming at providing support in the detection of 7-8-year-old children with reading disturbs. We proposed several metrics based on Information Retrieval measures for assessing the reading decoding capabilities. The alert thresholds have been empirically set and a preliminary validation on a little sample has been conducted. First results are encouraging: the app has detected two children seven years old with learning disorder on a little sample of children with problems of different age. The evaluation presented in this paper is very preliminary. To better validate the test there is the need of performing a controlled experiment with a large sample of children, classified as good reader and dyslexic. We are organizing this evaluation which will involve

Table 2: First validation data

ID	Effectiveness	Efficiency	Effficiency_bL	Effort	Age	Note
1	0.683	0.0683	0.256	8	9	autistic with borderline QI
2	0.952	0.087	0.346	11	10	slight mental retardation
3	**0.381**	**0.032**	**0.232**	12	10	language disorder and Dyslexia, II primary class
4	0.619	0.077	0.268	8	10	Dyspraxia and learning disturb, IV primary class
5	0.537	0.089	0.079	6	10	language and learning disorder
6	**0.476**	0.079	*0.079*	6	*9*	language and learning disorder
7	0.905	0.090	0.362	10	10	learning disorder
8	**0.450**	**0.045**	**0.045**	10	7	language and bilingualism disorder
9	**0.238**	**0.018**	**0.018**	13	7	language and learning disorder

many primary schools. If the positive results are confirmed we will deliver the app on the store for a massive use.

To obtain a deeper investigation, also the *Discrepancy* between skills in the specific domain (deficit in relation to expectations for the age and / or class attended) and general intelligence (appropriate for chronological age) may be evaluated [4]. We plan to add a general intelligence questionnaire to better identify children with cognitive problems w.r.t. dyslexic ones. The initial idea was to propose the tool as an online instrument for parents to provide a first suggestion. Thus, there is the need of updating the thresholds by taking the age into account. In particular, we need to verify the validity of the test for older children and, eventually, to set the threshold for each age. The app can also be extended to listening comprehension disturbs, by reading the sentence and letting the children compose it in the same way as the app at the present allows. Another idea may be to connect the app to a Learning Management System[7] to collect the data of an entire class e to compute the thresholds for the children belonging to it.

ACKNOWLEDGMENTS

The authors would like to thank all the experiment participants and Fondazione Salernum Anffas Onlus for its availability.

REFERENCES

[1] Alissa N. Antle, Min Fan, and Emily S. Cramer. 2015. PhonoBlocks: A Tangible System for Supporting Dyslexic Children Learning to Read. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '15)*. ACM, New York, NY, USA, 533–538. https://doi.org/10.1145/2677199.2687897

[2] V. R. Basili and H. D. Rombach. 1988. The TAME Project: Towards Improvement-Oriented Software Environments. *IEEE Transactions on Software Engineering* 14, 6 (1988), 758–773.

[3] Rosanna Cassino, Maurizio Tucci, Giuliana Vitiello, and Rita Francese. 2015. Empirical validation of an automatic usability evaluation method. *J. Vis. Lang. Comput.* 28 (2015), 1–22.

[4] AID Associazione Italiana Dislessia. 2007. Consensus Conference. Disturbi evolutivi specifici di apprendimento. Raccomandazioni per la pratica clinica definite con il metodo della Consensus Conference.

[5] Barbara Fioravanti Enrico Savelli, Stefano Franceschi. 2013. *SPEED Screening Prescolare EtÃă Evolutiva - DISLESSIA*. Erickson.

[6] Min Fan and Alissa N. Antle. 2015. Tactile Letters: A Tangible Tabletop with Texture Cues Supporting Alphabetic Learning for Dyslexic Children. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '15)*. ACM, New York, NY, USA, 673–678. https://doi.org/10.1145/2677199.2688806

[7] Rita Francese, Carmine Gravino, Michele Risi, Giuseppe Scanniello, and Genoveffa Tortora. 2015. Using Project-Based-Learning in a mobile application development course: An experience report. *Journal of Visual Languages & Computing* 31 (2015), 196 – 205.

[8] Uta Frith. 1985. Beneath the surface of developmental dyslexia. In *Surface dyslexia*. Colhearth and Marshall, 301–330.

[9] Ombretta Gaggi, Claudio Enrico Palazzi, Matteo Ciman, Giorgia Galiazzo, Sandro Franceschini, Milena Ruffino, Simone Gori, and Andrea Facoetti. 2017. Serious Games for Early Identification of Developmental Dyslexia. *Comput. Entertain.* 15, 2, Article 4 (April 2017), 24 pages. https://doi.org/10.1145/2629558

[10] Z Jamaludin, H Husni, and MM Alobaedy. 2018. In search for a viable pedagogical agent in assistive applications for dyslexic children. *Journal of Fundamental and Applied Sciences* 10, 6S (2018), 1758–1770.

[11] N. Juristo and A.M. Moreno. 2001. *Basics of Software Engineering Experimentation*. Kluwer Academic Publishers.

[12] B. Kitchenham, S. Pfleeger, L. Pickard, P. Jones, D. Hoaglin, K. El Emam, and J. Rosenberg. 2002. Preliminary Guidelines for Empirical Research in Software Engineering. *IEEE Transactions on Software Engineering* 28, 8 (2002), 721–734.

[13] D. LukeÅą. 2015. Dyslexia friendly reader: Prototype, designs, and exploratory study. In *2015 6th International Conference on Information, Intelligence, Systems and Applications (IISA)*. 1–6. https://doi.org/10.1109/IISA.2015.7388008

[14] Oppenheim. 1992. *Questionnaire Design, Interviewing & Attitude Measurement*. Pinter.

[15] Siti Khatijah Nor Abdul Rahim, Nur Hasni Nasrudin, Anis Zafirah Azmi, Rosida Ahmad Junid, Zainab Mohamed, Ini Imaina Binti Abdullah, et al. 2018. Designing Mobile Application for Dyslexia in Reading Disorder Problem. *International Journal of Academic Research in Business and Social Sciences* 8, 1 (2018), 628–646.

[16] F. Ricca, M. Di Penta, M. Torchiano, P. Tonella, and M. Ceccato. 2010. How Developers' Experience and Ability Influence Web Application Comprehension Tasks Supported by UML Stereotypes: A Series of Four Experiments. *IEEE Transactions on Software Engineering* 36, 1 (2010), 96–118.

[17] Filippo Ricca, Giuseppe Scanniello, Marco Torchiano, Gianna Reggio, and Egidio Astesiano. 2014. Assessing the Effect of Screen Mockups on the Comprehension of Functional Requirements. *ACM Trans. Softw. Eng. Methodol.* 24, 1 (Oct. 2014), 1:1–1:38.

[18] Richard M. Ryan and Edward L. Deci. 2000. Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology* 25, 1 (2000), 54–67. https://doi.org/10.1006/ceps.1999.1020

[19] Gerhard Schwabe and Christoph Göth. 2005. Mobile Learning with a Mobile Game: Design and Motivational Effects. *Journal of Computer Assisted Learning* 21, 3 (2005), 204–216. https://doi.org/10.1111/j.1365-2729.2005.00128.x

[20] Mariam R Sood, Annet Toornstra, Martin I Sereno, Mark Boland, Daniele Filaretti, and Anuj Sood. 2018. A Digital App to Aid Detection, Monitoring, and Management of Dyslexia in Young Children (DIMMAND): Protocol for a Digital Health and Education Solution. *JMIR research protocols* 7, 5 (2018).

[21] Morini I. Stella G., Scorza M. 2011. *Strumento per l'identificazione della lentezza nella lettura orale*. Giunti Scuola.

[22] P. A. D. Tore, S. D. Tore, L. A. Ludovico, and G. R. Mangione. 2014. MADRIGALE: A Multimedia Application for Dyslexia and Reading Improvement GAmifying Learning Experience. In *2014 International Conference on Intelligent Networking and Collaborative Systems*. 486–491. https://doi.org/10.1109/INCoS.2014.48

[23] Giuliana Vitiello, Rita Francese, Monica Sebillo, Genny Tortora, and Maurizio Tucci. 2017. UX-Requirements for Patient's Empowerment - The Case of Multiple Pharmacological Treatments: A Case Study of IT Support to Chronic Disease Management. In *IEEE 25th International Requirements Engineering Conference Workshops, RE 2017 Workshops, Lisbon, Portugal, September 4-8, 2017*. 139–145.

[24] C. Wohlin, P. Runeson, M. Höst, M.C. Ohlsson, B. Regnell, and A. Wesslén. 2012. *Experimentation in Software Engineering*. Springer.

[7] https://en.wikipedia.org/wiki/Learning_management_system