

## **zReader: a Mobile Game Suite for Improving Children's Reading Skills and Behavior**

**Fabício V. de A. Guerra<sup>1</sup>, José Amâncio M. Santos<sup>2</sup>,  
Dalton D. S. Guerrero<sup>3</sup>, Jorge C. A. de Figueiredo<sup>3</sup>**

<sup>1</sup>Programa de Pós-Graduação em Ciência da Computação  
da Universidade Federal de Campina Grande

<sup>2</sup>Departamento de Tecnologia da Universidade Estadual de Feira de Santana

<sup>3</sup>Departamento de Sistemas e Computação  
da Universidade Federal de Campina Grande

fabricio10@gmail.com, zeamancio@uefs.br,

{dalton, abrantest}@dsc.ufcg.edu.br

**Abstract.** *This work describes an empirical, educational and technological research conducted within the context of Brazilian formal literacy education problems. The main purpose of the research was to address the question of whether research-specific software for mobile devices could improve disadvantaged children reading related variables. In this regard, zReader, a mobile game suite for reading and storytelling, was developed and validated by means of an experiment with children attending Second and Third grades in a Brazilian public school. Results include positive inferential findings for children's reading skills improvement and book reading frequency.*

### **1. Introduction**

For many decades Brazilian literacy education problems have been a major concern among educational professionals and researchers, with illiteracy rates getting more and more concentrated and reflecting social-economical differences among Brazilian regions [Ferraro and Kreidlow 2004] – with notable literacy deficits in regions North and North-east [IBGE 2003].

These deficits were still an issue by the turn of the last century, with more than 15 millions of illiterates above the age of 15 [IBGE 2003] – and with more than 34 millions of *functional illiterates*, people who could read only basic sentences but still not enough to deal with day-to-day tasks at home or at work.

These numbers may be better understood by taking closer look at specific statistics about the early years of literacy education. Let us consider, for instance, the case of Caicó, a typical Brazilian northeastern city. Data from 2014 indicated that about 2/3 its 3<sup>rd</sup> grade children – that just completed the so called *alphabetizing cycle*<sup>1</sup> – could not locate an explicit information in a text segment and could not infer cause/effect relation in a joke or

---

<sup>1</sup>Brazilian alphabetizing cycle consists of the first 3 elementary grades, where children begin to have formal reading education. Children from 2<sup>nd</sup> grade on are supposed to have the minimum reading skills required to participate on this experiment

fable [INEP 2017a]. Data from 2015 [INEP 2017b], in addition, showed that after 9 years of formal public education, not a single student reached the highest reading level for the test. That is quite a discouraging picture for parents, researchers and educators.

While our children struggled with public literacy education, mobile computer devices became populars and low-cost. Several experiments were conducted around the world to test how technology could be properly used in educational environments, including, of course, literacy education/reading comprehension (both terms will be used interchangeably along the text). However, for Brazilian Portuguese literacy specific solutions, not too many technological options are available, as explicitly stated in a recent Brazilian Portuguese specific systematic mapping [Gaspar et al. 2015] and as one may infer from a broader systematic review [Takacs et al. 2015] that selected no work for that language. Moreover, two important remarks are presented in the later research. First, it shows that features like game playing are, for many times, harmful for reading comprehension. Second, it shows that non-interactive multimedia features favor low socio-economic family children (“disadvantaged children”) for reading skills, working on their linguistic comprehension deficiencies.

Under all these circumstances, this work focused on using highlights from researches concerning this later educational software application remarks in the production of serious games to address issues related to the former discussion about poor literacy rates at public school low-income children in Brazil. The research was targeted at 2<sup>nd</sup> and 3<sup>rd</sup> grade elementary school children. An educational mobile game technology – zReader – was proposed and empirically tested in the context of an experiment in a public low-income neighborhood school. Inferential results favored experimental group over control group for reading skills improvement and also for general book reading frequency. The experiment addressed the following research questions.

**RQ1** - Does using zReader make children read more books, compared to children not using zReader?

**RQ2** - Does using zReader improve children’s reading skills better, compared to children not using zReader?

## 2. The zReader and its theoretical foundation

zReader is a mobile application consisting of several implementations of literacy games. The games may be implemented according to three different designs and, in this paper, we present two of these, the ones that were created specifically to impact on 2<sup>nd</sup> and 3<sup>rd</sup> grade children’s literacy skills and reading behavior. The designs involve user interface structure and interactions guidelines that coordinate the activities of reading a narrative text, building a corresponding cartoon animation and dubbing the cartoon.

The backdrop educational framework considered for the application design was Papert’s Constructionism [Papert and Harel 1991] which, on its turn, was developed on Piaget’s Constructivism [Müller et al. 2009]. In brief, as Constructivism more generally stresses on subject *interaction* with the object of study as a mean to achieve learning, Constructionism directs this interaction to the *creation* of artifacts of any kind (in our case, an animated cartoon), in order to promote learning “in a context where the learner is consciously engaged in constructing a public entity” [Papert and Harel 1991].

The mobile application also includes gamification [Groh 2012] and Serious Game [Ratan and Ritterfeld 2009] perspectives, since it uses “game elements for serious purposes” [Deterding et al. 2011]. Each type of game may be regarded as a level and the interactions always take place with a literacy-related purpose disguised as game goal or task.

Specific design decisions were based on cognitive theories that focus on brain logic architecture and how instructional material may foster learning by respecting architecture characteristics and limitations. These theories are Sweller’s Cognitive Load Theory (CLT) [Sweller 1994], Mayer’s Cognitive Theory of Multimedia Learning (CTML) [Mayer 2005] and Schnotz’s Integrated Model of Text and Picture Comprehension (ITPC) [Schnotz 2005]. CTML, a learning centered theory, is based on CLT, which is more general theory of human cognition [Mayer and Moreno 2010]. ITPC, on its turn, builds upon CTML structural components to explain not only multimedia learning but single-medium learning [Horz and Schnotz 2010]. It is according the these theories and some of their common concepts like limited working memory and generative processes that the games have a design without complex interactions or fancy features that may take children’s focus away from the inner narrative stories. Building the cartoons from scratch was also avoided. The whole user interactions are based on discrete touches, drags, drops, and simple gestures like “page turns”.

## 2.1. Text/Video Translation

The first game design of zReader is named “Text/Video Translation” (TVid, for short). In general, a TVid game consists of reading pages of a narrative fictional story and choosing proper cartoon animation for visually representing the text page read. When the animation for the last page is chosen, the game moves to a “Record Room” (RR, for short) where the chosen sequence is voiced over.

An example user interface for a particular TVid implementation is depicted in Figure 1(a). The interactions take place as follows: below the text component, two unlabeled buttons can trigger different animations in the video component at the right side of the screen. The goal is to watch both animations, choosing the one that better suits what is written in the text component (the page). The main idea behind the two animation options for a given page is to adapt CTML’s reflection principle [Moreno and Mayer 2007] – in this case, reflection is induced by game interaction. When the check button is touched, the next text page, with corresponding video animations, are automatically loaded. Page turn gestures – left-to-right and right-to-left lines – may be used to navigate forth and backwards in the text.

Figure 1(b) depicts a particular RR implementation. It includes a play button, labeled with a page number, a record button and a play-all button, along with a video component. The animation segments to be recorded reflects children’s choices in TVid and may be different from the written narrative story. Wrong choices lead to senseless final cartoons (“bad cartoons”, as children name them), so the children have to pay attention to the text when choosing visual representations. The recording, or the final cartoon, in this case, may act as a reward to the playing children. If they play TVid right, their cartoon will be a “good” cartoon. Again, page turn gestures are used for navigation.

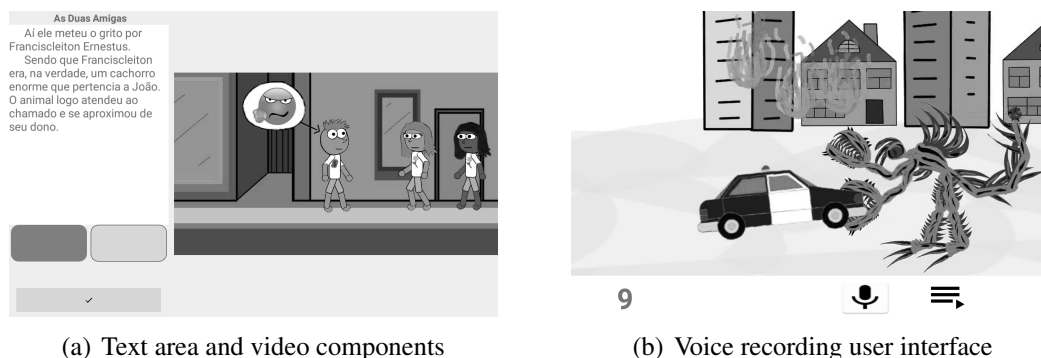


Figure 1. TVid user interface



## 2.2. Concept Puzzle

The second game design of zReader is named Concept Puzzle (CPuz, for short). In general, CPuz games consist of reading a narrative text and then arranging labeled puzzle pieces in sequence – each piece representing one of the text pages just read. The pieces are labeled with words that are central to the corresponding text page comprehension. We adopted the term *Concept Puzzle* instead of *Word Puzzle* to stress the important semantic relation between text segment content and puzzle piece label.

Figure 2 depicts the user interface for the *puzzle screen* of a particular CPuz implementation. There is a right *concept* panel containing labeled buttons, a bottom *montage* panel initially unlabeled and a video component for cartoon animation playing. In addition, there is a *text icon* in the bottom-right side of the screen (for reading the text).

The goal of CPuz games is to arrange all the concepts together by dragging them into the montage panel in a sequence that reflects the storyline of the underlying narrative text. To help with this task, the game plays a corresponding animation when a concept is touched. After the whole set of animation pieces is assembled together, the game moves to a RR screen for voice recording. Children may go back to the puzzle screen if they realize, while recording the cartoon, that they have made wrong choices.

CPuz design clearly imposes more cognitive load on children's working memory by the nature of its interactions and the number of options involved. It is important to note, though, that CPuz is designed for children with good reading comprehension skills that might be bored to play TVid games.

### 3. Materials and Method

A sample of  $N=77$  volunteer children was available for the experiment (31 from 2<sup>nd</sup> grade and 46 from 3<sup>rd</sup> grade), all of them with proper parental consent. They studied at the same public school in a low socio-economic neighborhood at the city of Caicó(RN), Brazil. Also, all children were at a regular situation with regards to age/grade relation and had no diagnosed cognitive problems. A control group of children attending regular computer lab classes (conducted by school teachers) was compared to an experimental group of children attending *optional* zReader sessions. The children were randomly assigned to control or experimental group.

Two dependent variables were observed: **book reading count** and **reading skills improvement**. The assessment of the **book reading count** variable was performed by means of a school reading program for children in 3<sup>rd</sup> grade ( $N = 46$ ) – in the program, all the books they read were counted. By program rules, children who read more books are granted prizes and medals each 2 and a half months, so we were quite confident that they did present to school staff any book they came to read. Reading comprehension was assessed by means of quantitative standard reading test TCLPP [Seabra and Capovilla 2010], before and after the zReader sessions. The difference between post and pre-test scores stands for **reading skills improvement**. We summarize the experimental setup in Table 1, showing dependent variables, their values/assessment and sample size descriptors.

**Table 1. A summary of the experimental setup**

Dependent variable	Sample size		Value	
	Cont	Exp	Initial	Final
book reading (3 <sup>rd</sup> grade only)	24	22	0	number of books read
reading comprehension	39	38	TCLPP (pre-test)	TCLPP (post-test)

#### 3.1. Treatment dynamics

Two training sessions and eight regular sessions, one session a week, for each grade, were scheduled with experimental group children. Training sessions were collective and consisted of one member from the research team explaining the children how to interact with zReader. Regular sessions consisted of children themselves interacting with zReader in all of its features with on-demand help by research team. The games were played in pairs and, in training sessions, children were explained the “rules” for playing them as well as the rules for playing them in pairs, as we now describe the later.

In TVid, both children read the text part and one of them was responsible for playing the animations and suggesting the right one. If the other child agreed on the choice, they would confirm it and go to the next page. If he/she did not agree on the choice, they both had to discuss their reasons and decide together on which choice was best related to the text. The roles were inverted for each “page turn”. In CPuz, the dynamics were quite the same – the children just had more options of animations for choosing. For recording the cartoon, the children were free to choose the characters they would play and any other possible action about recording when there was no balloon dialogue visible (in which case they had to read the balloon).

It was allowed for children who missed sessions, to make extra sessions to replace the missing ones just as control group children could do with regular computer lab classes. They were also allowed to read/record another story if they wanted, provided that they had finished the current story.

A final and important remark on treatment dynamics is that experimental group children began the treatment with TVid games and passed to CPuz games when they found TVid games too “easy” for them. It was their decision to move to a *harder* game (a harder “level”, as they name it).

### 3.2. Treatment notes

A total of 16 game implementations were used along the way, all of them implemented as Android 5.0.2 projects and installed on 5 Samsung Galaxy Tab 4 tablets with quad-core 1.2GHz processor and 10.1 inches 1280 x 800 pixels screen size.

All the texts used in the implementations, in terms of sentence complexity, were chosen with regards to what the children were *probably* able to read. The implementations were not to introduce Portuguese language topics or replace teacher classes, but to develop abilities on the top of what children knew. Obviously, it was not known beforehand what were exactly the words and syntactic clause structures that they could understand, but children’s text books were analyzed and their teacher was consulted with that in mind, so the narrative texts could be accordingly adapted. Eventual replications or adaptations of the experiment must take that into consideration.

### 3.3. Analysis procedure

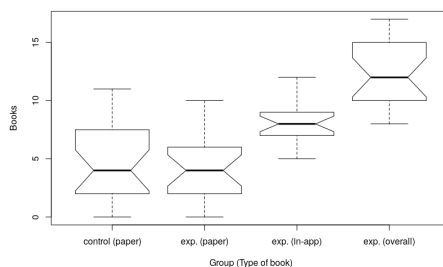
The analysis of both **book reading count** and **reading skills improvement** variables were based on the following approach. Initially, the dispersion of the variables was visually observed in notched box plots. Then, depending on the visual clues, we proceeded to distribution checking and inferential comparisons between groups. Shapiro-Wilk test was used to check for normality in distributions. Parametric Welch two sample t-tests were used to investigate inferential differences between normally distributed data sets and non-parametric Wilcoxon test was used whenever any of the considered data sets was not normally distributed. When differences were found, the effect sizes were calculated either by Cohen’s *d* (for normal cases) or Cliff’s *delta* (non-normal cases).

## 4. The results or how data analysis answers research questions

### 4.1. Book reading count

The analysis of the book reading count variable considered only the 3<sup>rd</sup> grade students (no corresponding data was available for 2<sup>nd</sup> grade). The hypothesis is that the zReader positively impacts the number of books the students read. The box plots in Figure 3 show distributions of **book count** considering control and experimental groups for paper and in-app book counts.

It is quite clear, by visually comparing the box plots for **exp. (overall)** and **control (paper)**, that experimental group children read more books than control group children. The same occurs by comparing the number of books read in zReader with the paper books read by control group children – **exp. (in-app) versus control (paper)**. As for



**Figure 3. Distribution of the book reading count variable**

paper book reading alone, the behavior seems uniform along control and experimental groups – **exp. (paper) versus control (paper)**. Finally, it is also possible to note that children in experimental group also read more in-app books as compared to themselves - **exp. (in-app) versus exp. (paper)**. These observations evidence that, with zReader, children read a lot more overall books without affecting paper reading behavior.

Table 2 summarizes null hypothesis, statistical tests and effect sizes in order to confirm the results visually observed in Figure 3. The tests confirm previous analysis indicating that zReader positively impacts on the number of overall books that children read (**exp. (overall) = control (paper)**). The magnitude of this impact is measured as Cliff’s delta and indicates a high effect size value of  $d = 0.88$ . The test also indicated a 92.61% chance that a child randomly chosen from experimental group read more books than a randomly chosen case from control group.

These results answer **RQ1**, that is, **using zReader makes children – 3<sup>rd</sup> grade children, in this case – read more books.**

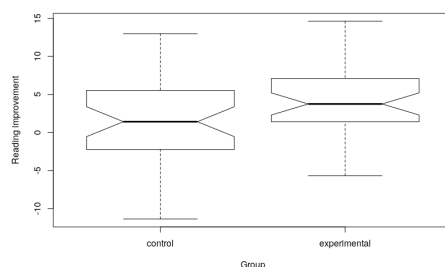
**Table 2. A summary of the statistical tests**

Null hypothesis	Normality test p-value	Hypothesis test p-value		Cliff’s delta
		Welch	Wilcox	
exp. (overall) = control (paper)	<b>0.001</b> - 0.14	–	<b>&lt;0.0001</b>	0.88
exp. (in-app) = control (paper)	<b>0.01</b> - 0.14	–	<b>0.0002</b>	0.6
exp. (in-app) = exp. (paper)	<b>0.01</b> - 0.06	–	<b>0.0001</b>	0.7
exp. (paper) = control (paper)	0.06 - 0.14	0.59	–	–

#### 4.2. Reading skills improvement

For reading skills improvement, our sample involves children in both 2<sup>nd</sup> and 3<sup>rd</sup> grades. A positive improvement was expected for both groups, once all children were receiving daily instructions on reading by means of conventional classes. Our hypothesis is that zReader improves children reading skills better for the experimental group. This scenario is depicted in Figure 4, where notched box plots for reading improvement in control and experimental groups are shown. Visually, we can see that experimental group median is above control group confidence interval (denoted by the notches). Thus, we proceeded with inferential tests.

Normality was assumed for the distributions of both control and experimental groups ( $p > 0.7$  and  $p > 0.3$  in Shapiro-Wilk test). The null hypothesis was that the



**Figure 4. Box plots for reading skills improvement variable**

groups were equal and a one-tailed Welch two sample t-test was significant ( $p < 0.014$ ). Thus, we accept the alternative hypothesis that mean reading improvement in experimental group is greater than in control group. Furthermore, a moderate effect size of  $d=0.51$  (Cohen's  $d$ ) was found. These results answer **RQ2**, that is, **using zReader improves children's reading skills better as compared to children not using zReader**.

## 5. Discussion or how these results are any good

The most important point to be discussed is whether zReader directly caused the observed better reading skills improvement or whether it caused more reading practice that, on its turn, caused the better reading skills improvement. It is maybe counter-intuitive, but even daily contact with conventional books and improved taste for reading are not always observed as causing literacy skills improvement [Goux et al. 2016], so we should not despise the effects of game dynamics in this experiment. As for the moderate effect size, getting greater values with technology alone, without counting on traits like teacher training or daily exposure [Song et al. 2014, Santiago et al. 2016] does not seem to be a simple task. Moreover, conclusions from a related systematic review [Takacs et al. 2015] report "small but significant addition" over specific skills like story comprehension.

The behavior of reading more books is also important for alphabetizing children. Specially because this behavior is possibly associated with taste for reading, which is cited by 5-10 years old children as the most important factor for motivating them to read books, as published in last Pro-Livro Institute report [Failla 2012, p. 193]. Thus, it is possible that zReader affected this important variable and that the high number of overall reading is a reflex of that. One may argue that is logical to have more overall book reading, since experimental group children were given access to digital devices, but it is important to remember that the zReader sessions were optional. This is also counter-intuitive, but it has been observed that access to eReading enabled devices may be associated with *less* reading in general [Merga and Roni 2017]. In sum, when it comes to reading and technology, there seems to be no logical path to follow. The best way to draw conclusions is by running controlled experiments.

Lastly, we also have to consider the developed application with regards to its possible integration into formal education environments. There are great published solutions, like TaBooGa [Linke et al. 2017] (a hybrid e-book connected to external tangible elements), that also work on literacy and motivation but are, in terms of costs, very far away from being implemented in Brazilian public schools. Mobility features of zReader make possible for a relatively low cost mobile lab with a few tablets to serve several different



schools. Moreover, a positive evidence that children had a good user experience with the software was the fact that they attended almost all zReader optional sessions they had the opportunity to attend – 3 refusals out of something like 300 game plays. These considerations address two essential perspectives for integration: implantation costs and children’s user experience with the application.

## 6. Conclusions, limitations and what to do next

This research has some limitations when it comes to the broader sense of Brazilian literacy problems, for it was not feasible to experiment over a random sample from such a huge population as Brazilian disadvantaged children. So our approach was the *random assignment* of a smaller group of volunteers to the treatment, as most researchers’ approaches are. Replication of the experiment with other groups may compensate the lack of a proper population random sample.

Currently, we are developing and validating zReader games targeted at 1<sup>st</sup> graders and designing experiments in order to find evidence about biological reasons that may have caused the results we have found for 2<sup>nd</sup> and 3<sup>rd</sup> graders. The main possibility under investigation is that game dynamics improve children’s attention into the text parts of the game.

Finally, it is important to note that we’ve been conducting experiments with children in alphabetizing cycle since 2015. From unsuccessful solutions, reworked design decisions and pilot experiences to the current working version of zReader, we dealt with something like 400 alphabetizing children in 5 private and public schools of 3 different cities. Even though it is not possible to make such a claim from a scientific perspective, we find very, very unlikely that one may succeed in getting these children to read paper or conventional digital books just by offering them the possibility to do so once a week. Literacy would hardly be an issue if this simpler approach was effective and it was effective for reading with our application. The extremely low attendance refusal to the optional in-app book reading was, in that sense, if not a formal result, a very encouraging observation.

## References

- Deterding, S., Khaled, R. N., and Nacke, L. (2011). Gamification: Toward a definition. In *CHI 2011 Gamification Workshop Proceedings*.
- Failla, Z. (2012). *Retratos da leitura no Brasil 4*. Imprensa Oficial do Estado de São Paulo.
- Ferraro, A. R. and Kreidlow, D. (2004). Analfabetismo no brasil: configuração e gênese das desigualdades regionais. *Educação & Realidade*, 29(2).
- Gaspar, W., Oliveira, E., and Oliveira, K. (2015). Aprendizagem da língua portuguesa com dispositivos móveis: Um mapeamento sistemático da literatura. In *SBIE*, volume 26, page 140.
- Goux, D., Gurgand, M., and Maurin, E. (2016). Reading enjoyment and reading skills: Lessons from an experiment with first grade children. *Labour Economics*.
- Groh, F. (2012). Gamification: State of the art definition and utilization. *Institute of Media Informatics Ulm University*, 39.

- Horz, H. and Schnotz, W. (2010). Cognitive load in learning with multiple representations. In *Cognitive load theory*, pages 229–252. Cambridge University Press New York, NY.
- IBGE, C. D. (2003). Mapa do analfabetismo no brasil. *Brasília, MEC/INEP*.
- INEP (2017a). Ana - sistema. <http://ana.inep.gov.br/ANA/>. Accessed: 2017-03-07.
- INEP (2017b). Distribuição dos alunos por nível de proficiência. <http://www.qedu.org.br/cidade/5409-caico/proficiencia>. Accessed: 2017-03-07.
- Linke, R., Kothe, T., and Alt, F. (2017). Tabooga: A hybrid learning app to support children's reading motivation. In *Proceedings of the 2017 Conference on Interaction Design and Children*, pages 278–285. ACM.
- Mayer, R. E. (2005). Cognitive theory of multimedia learning. In *The Cambridge handbook of multimedia learning*. Cambridge university press.
- Mayer, R. E. and Moreno, R. (2010). Techniques that reduce extraneous cognitive load and manage intrinsic cognitive load during multimedia learning. In *Cognitive load theory*, volume 132. Cambridge University Press New York, NY.
- Merga, M. K. and Roni, S. M. (2017). The influence of access to ereaders, computers and mobile phones on children's book reading frequency. *Computers & Education*.
- Moreno, R. and Mayer, R. (2007). Interactive multimodal learning environments. *Educational Psychology Review*, 19(3):309–326.
- Müller, U., Carpendale, J. I., and Smith, L. (2009). Introduction: Overview. the context of piaget's theory. In *The Cambridge companion to Piaget*, pages 1–44. Cambridge University Press.
- Papert, S. and Harel, I. (1991). Situating constructionism. In *Constructionism*, volume 36, pages 1–11.
- Ratan, R. and Ritterfeld, U. (2009). Classifying serious games. *Serious games: Mechanisms and effects*, pages 10–24.
- Santiago, L., Dionisio, M., and Sampaio, R. (2016). O uso das tecnologias digitais na busca da superação do analfabetismo. In *CBIE*, volume 5, page 800.
- Schnotz, W. (2005). An integrated model of text and picture comprehension. In *The Cambridge handbook of multimedia learning*. Cambridge university press.
- Seabra, A. G. and Capovilla, F. C. (2010). Teste de competência de leitura de palavras e pseudopalavras (telpp). *São Paulo, SP: Memnon*.
- Song, H. S., Pusic, M., Nick, M. W., Sarpel, U., Plass, J. L., and Kalet, A. L. (2014). The cognitive impact of interactive design features for learning complex materials in medical education. *Computers & education*, 71:198–205.
- Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and instruction*, 4(4):295–312.
- Takacs, Z. K., Swart, E. K., and Bus, A. G. (2015). Benefits and pitfalls of multimedia and interactive features in technology-enhanced storybooks a meta-analysis. *Review of educational research*.