Towards a digital syllable-based reading intervention: An interview study with second-graders

Philipp Schaper¹, Melissa Donnermann¹, Nicole Doser¹, Martina Lein¹, Anna Riedmann¹, Sophia C. Steinhaeusser¹, Panagiotis Karageorgos², Bettina Müller², Tobias Richter², Birgit Lugrin¹
Human-Computer-Interaction¹ and Educational Psychology²
Julius-Maximilians-University of Wuerzburg
Wuerzburg, Germany
{philipp.schaper, anna.riedmann, bettina.mueller, birgit.lugrin}@uni-wuerzburg.de

Reading is an essential ability and a cornerstone of education. However, learning to read can be challenging for children. To scaffold young learners, a number of reading interventions were developed, including a syllable-based approach in German, which has proven to be successful, but resource and time consuming through individual interaction by educators. To improve the reach of the reading intervention, we present the first step towards a digital intervention, following a human-centred design approach. In this contribution, we present the implementation of a digital prototype, developed with the feedback of expert evaluations, as well as an interview study with second-graders. The results of interviews with children showed that the app is suitable to be applied in the target age group, that children had fun using it and were motivated to further do so. The study also provides design implications for transferring an analogue concept into a digital application.

1. INTRODUCTION

Around 6.2 million people in Germany cannot read properly (Grotlüschen & Riekmann, 2011). The origin of the inability or a deficiency in reading capabilities can often be traced back to the initial attempts to learn reading during primary school. This problem is reflected in around 15% of German fourth graders, who show deficits in extracting meaning from presented texts (Bos, Tarelli, Bremerich-Vos, & Schwippert, 2012). In the overall reading performance, German children showed a lower mean performance score relative to most other European countries (Hußmann et al., 2017). This is due a stagnating overall score for Germany, which has not improved significantly since 2001 in contrast to many other European countries. However, there is also high variability in the reading abilities, with an increase in children on the lowest level of competence to 6% from 3% in 2001.

Given the high transparency of the German language that allows reading words on a letter-by-letter basis, even poor readers can read words accurately. However, this process is resource demanding and prone to errors (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001), and thus not suited for understanding the meaning of a text. To read words accurately and fast, readers need to learn to use units larger than single letters (i.e., syllables and morphemes) – a step most poor readers in German experience difficulties with. Due to the relevance of reading abilities for knowledge acquisition (Nagler, Lindberg, & Hasselhorn, 2017) as well as social and cultural activities (Naumann, Arlett, Schneider, & Stanat, 2010), problems in reading acquisition should be addressed as early and as widespread as possible.

A number of analogue interventions were designed for poor readers in primary school in different languages (e.g., Ecalle, Magnan, & Calmus, 2009; Heikkilä, Aro, Närhi, Westerholm, & Ahonen, 2013), including German (e.g., Müller, Richter, & Karageorgos, 2020). The latter is a syllable-based reading intervention to foster word reading skills for groups of second-graders. A digital version thereof could facilitate learning anytime anywhere, allow personalized help and potentially provide a fun approach of additional training. In this contribution, we first present a theoretical overview on reading acquisition and the reading intervention. Second, we developed a high-fidelity prototype that digitalizes a fraction of the syllable-based reading intervention, alongside with an interview study with second-
graders that aims to explore whether a digital version could be employed with the target age group. Our contribution is twofold: (1) We demonstrate a systematic approach to transferring established concepts into a digital format by adapting a validated analogous training for mobile devices; (2) We derive relevant design implications based on a target group oriented evaluation of the application.

1.1. Theoretical background on reading acquisition

In general, the acquisition of reading skills on the word-, sentence- and text-level of young children varies within the development process. Children quickly improve their reading abilities during the early grades, whereas the degree of improvement is lower in third-graders and subsequent grades (Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998; Logan et al., 2013). Due to the trajectory of reading skills development during grade one through four (Foorman et al., 1998), early intervention is essential to train deficient reading abilities with the aim to return children to the path of normal reading development (Francis, Shaywitz, Stuebing, Shaywitz, & Fletcher, 1996; Müller, Otterbein-Gutsche, & Richter, 2018). Without intervention, children, who have poor reading skills in early grades, are highly likely to remain poor readers (Landerl & Wimmer, 2008).

As children progress in reading development, they should reach the consolidated phase as they are able to memorize a larger amount of sight words by extracting larger chunks of grapheme-phoneme connections, including syllables (Ehri, 1995, 2005). The more words can be read by orthographic representations like syllables the faster word recognition is possible. Because syllables are often used in multiple words, readers might profit from transfer effects. Syllable-based reading therefore qualifies as a training technique to gain knowledge which is transferable to unknown texts or words (Huemer, Aro, Landerl, & Lyytinen, 2010). However, poor readers in German often experience difficulties in using (sub)lexical units like syllables or morphemes to recognize words faster via orthographic decoding processes (Landerl & Wimmer, 2008). In consequence, their word recognition is slow and error prone which leads to difficulties in sentence- and text-based reading comprehension. Their trajectory of reading skill development is less steep compared to readers with good word recognition skills (Pfoist, Hattie, Dörfler, & Ärtelt, 2014).

Using the capacity of syllables, Müller et al. (2018; 2020) developed a syllable-based reading intervention for second-graders. The intervention aims to enhance the word recognition processes by repeatedly reading and segmenting syllables using 24 different game-like training sessions. The word material within the manualized intervention was selected based on the 500 most frequent German syllables in texts typically read by 6-8 year old children (cf. database childLex, Schroeder, Würzner, Heister, Geyken, & Kliegl, 2015). The intervention aims to strengthen the mental representations of syllables and words consisting of these syllables to foster the accuracy and fluency of word recognition. Thus, the intervention should support poor readers to reach the consolidated phase instead of reading words letter-by-letter.

Müller et al. (2018; 2020) evaluated the reading intervention in the form of biweekly training sessions organized by specifically prepared trainers in addition to regular lessons. Participating children showed significant improvements regarding phonological and orthographic word recognition, hence being able to read words more quickly and accurately compared to same-skilled poor readers in the control condition. Participants also profited from enhanced reading comprehension due to transfer effects (Müller, Richter, & Karageorgos, 2020; for results in Grade 4 see Müller, Richter, Karageorgos, Krawietz, & Ennemoser, 2017). The results of the intervention highlight the value of early support in reading development to effectively improve reading skills (Klatte, Steinbrink, Bergström, & Lachmann, 2013; Tacke, 2005). However, conducting such an intervention in small groups is time and resource consuming. Additionally, even though group-based interventions have been shown to work, individual support for each learner would be beneficial (Groth, Hasko, Bruder, Kunze, & Schulte-Körne, 2013).

1.2. Related work on the development of app-based learning

Mobile apps have been introduced into the field of education for young learners for a wide range of topics ranging from science to language acquisition (e.g., Henderson & Yeow, 2012; Walter-Laager et al., 2017). Several studies have highlighted the positive effects of integrating tablet apps in educational contexts (e.g., Bastian & Aufenanger, 2017; Schoppek & Tulis, 2010), even compared to desktop applications (Sung, Chang, & Liu, 2016). There are also indications that mobile apps increase motivation (Su & Cheng, 2015; Tillmann & Bremer, 2017). Further advantages include access at any time, versatility, adaptability (Rossing, Miller, Cecil, & Stamper, 2012) and individual feedback to consolidate learning progress (Blok, Oostdam, Otter, & Overmaat, 2002).

Although a great amount of educational apps exists for young learners, quantity does not ensure quality of content (Papadakis, Kalogiannakis, & Zaranis, 2018). Haßler, Major, and Hennessy (2016) criticize that only a small subset of tablet apps is empirically tested, and that only a few studies adhere to
minimum quality criteria, stressing the need for further empirical research. The general notion that adequately tested mobile apps might benefit learning and that introducing technology into the classroom can additionally enhance interest and motivation appears to be well supported (Hochberg, Kuhn, & Müller, 2018).

There are several apps with the overarching goal to practice and improve reading. For example, there are apps with picture books for beginners in reading (e.g., Cahill & McGill-Franzen, 2013), or multimodal e-books to train children’s reading skills (e.g., Morgan, 2013), indicating different possibilities to support reading acquisition. Focusing on spelling in German, there are similar issues with the apps in terms of quality (Fleischhauer, Schledzewski, & Grosche, 2017). To date, no evidence-based app using a syllable-based approach with focus on orthographic and phonological processes is available.

The contrast between the large amount of reading apps available and the lack of using established and evaluated approaches highlights a need for a scientifically driven development process (Hirsh-Pasek et al., 2015). In order to obtain the benefits of mobile learning, the app needs to be designed focusing on its usability and handling by children. Comfort and ease of use are also important aspects.

When adapting an established reading intervention from an analogous course-based format to a mobile app, several aspects have to be considered. For example, motivational aspects and self-directed learning become more relevant, because users interact with the app by themselves, and potentially have no external guidance (Ericsson, Krampe, & Tesch-Römer, 1993). A mobile app should promote self-directed learning on different levels of difficulties in pupils, similar to the benefits reported for technology-rich learning environments (Fahnoe & Mishra, 2013; Rashid & Asghar, 2016; Rossing et al., 2012; Underwood, Luckin, & Winters, 2012).

1.3. Contribution

Acquiring recognition and understanding of written words is one of the main educational goals in primary school. For poor readers learning to read in German, a syllable-based intervention has been shown to positively influence word reading skills and, in consequence, text-based reading comprehension (Müller, Richter, & Karageorgos, 2020; Müller et al., 2017). However, the analogue application of this approach is time and resource consuming, as many of the conventional interventions highly rely on individual interaction by educators to convey the learning material. A digital version of this intervention for reading acquisition could potentially overcome these limitations. However, the transfer of such an evidence-based face-to-face intervention into a digital version should be conducted with a human-centred design approach. As noted by Görgen, Huemer, Schulte-Körne, and Moll (2020), until recently the focus for development for children with reading problems has been on analogue trainings with experts. To the best of our knowledge, even though steps towards digital reading trainings have been made, they are not set up as transferred version of a tested analogue training. Even though steps in this direction have been taken by comparing pen-and-paper and digital content in math education (Maertens, Vandewaetere, Cornillie, & Desmet, 2014).

As a first step, this study addresses to which degree such analogue teaching material, which is originally presented by an instructor, can be transferred into a digital version. Perceived usability and enjoyment regarding the learning material might be crucial in determining the frequency of use of the digital version, both by children at home and teachers in their classes. Only if these potential motivational obstacles are overcome and the underlying principles of the intervention can be conveyed, a subsequent larger investigation on the actual learning outcome in the target group can be considered.

We therefore present the first step of a human-centred design approach to implement a digital version of the reading intervention by Müller, Richter, and Otterbein-Gutsche (2020), providing an operational high-fidelity prototype that integrates expert feedback from educational psychologists and primary school teachers as well as feedback from second-graders, the target group of the reading intervention. Our work demonstrates a systematic approach to transferring an analogue concept into a digital application and design implications derived from this process.

It is important to keep in mind that the current state of the app is not yet a complete digital version of the original intervention, for which similarly positive results on learning outcomes may be expected. However, the prototype is intended as an elaborated proof of concept on whether the intervention has the potential to be developed into a mobile, easy-to-use and motivating application, which is accepted by pupils and teachers alike and able to improve reading abilities, based on empirically tested scientific material. To this end, we addressed the following research questions:

RQ1: Can the central elements of the analogue syllable-based reading intervention be successfully transferred into a digital version?

RQ2: Are second-graders motivated to use the app again?

RQ3: Do the second-graders enjoy using the app?
2. PROTOTYPE OF THE READING APP

For the prototype, three educational games used in the reading intervention (Müller, Richter, & Otterbein-Gutsche, 2020) were chosen for implementation in Android Studio (Google, 2019). First, a selection of games were implemented as low-fidelity prototypes using Axure (Axure Software Solutions Inc., 2019). Second, three of these prototypes were chosen for implementation as an Android app, based on their ability to capture essential aspects of the reading intervention such as the ability to separate words into syllables or the acquisition of syllable rhythm and syllable classification.

2.1. Analysis of requirements

By digitalizing the complete syllable-based intervention, it could potentially benefit from the advantages of technology-based tools (Biancarosa & Griffiths, 2012; Sung et al., 2016). A digital version could also avoid the logistic drawbacks of the analogue intervention, which requires specifically prepared teachers. Furthermore, as the app is used by each child individually, the pace for the training is not affected by the group (Groth et al., 2013) and an adaptive training at the child’s individual reading level is possible.

We identify four main criteria for our app based on relevant requirements for mobile education applications for children as proposed by Mkpojiogu, Hussain, and Hassan (2018) and the PACMAD (People At the Centre of Mobile Application Development) usability model by Harrison, Flood, and Duce (2013): (1) efficiency in terms of adequate and accurate feedback, (2) effectiveness with regard to the comprehensibility of content and navigation elements, (3) the self-descriptiveness of the individual games and the app overall (learnability) and (4) child-friendly design to assure user satisfaction and reduce cognitive load. Additionally, we focused on the correctness of content. These requirements have to be considered for the adaption of the analogue version to a digital version, which is closely based on the original training while being extended by meaningful modifications, such as an immediate feedback mechanism, to fully exploit the potential of a digital adaptation. For subsequent development, further requirements such as adapting the app’s overall difficulty level to the children’s individual skills requires further analysis of user data, e.g. by collecting performance data.

The correctness of the content is addressed by a direct transfer from the analogue reading intervention to the digital version, and a feedback loop via an expert evaluation. The four requirements for educational apps for children (Mkpojiogu et al., 2018) are dependent on design choices during the development process. The crucial idea behind these requirements is that second-graders, especially with reading deficiency, should be able to successfully interact with the app. Therefore, vivid explanatory videos that address both the navigation within the app, as well as the content and procedures within each individual game are integrated. Additionally, visual and acoustic feedback are used throughout the interaction using simple, consistent sounds with the main purpose to indicate if a given solution to a task was right or wrong (Mejtoft, Lindberg, Söderström, & Mårell-Olsson, 2017). By implementing colorful and expressive graphics, the app is designed for the target age group of the reading intervention.

Finally, because scientifically developed technologies are often not integrated into everyday classrooms (Scherer, Tondeur, Siddiq, & Baran, 2018), the to be developed app had to be accepted not only by experts, but also by pupils as well as their teachers to be of practical use.

2.2. Main concept of the digital intervention

The main menu of the app is depicted as a map of an island structured in different areas represented by distinctive symbols corresponding to the three respective games (flowers, lighthouse, and school). After selecting one of the games, the user can choose between three difficulty levels. Completing a game in one of the difficulties returns the user to the difficulty selection screen.

Within each difficulty level, five tasks are presented sequentially, with a progress bar illustrating the amount of finished tasks per level. Each task is set to be worked on in an answer-until-correct style (Blair, 2013; Epstein, Epstein, & Brosvic, 2001), allowing the learner to reach the correct solution eventually. Users are also allowed to abandon the currently played level and return to the difficulty selection screen.

The design of the buttons was kept simple and comprehensible. The difficulty selection screen for each respective game as well as the levels themselves use an image of the previous app layer in the upper left corner as a button to return to this layer. To ensure proper understanding regarding the functionality of the app, tutorial videos are displayed when starting the app for the first time and before each game. Each tutorial video can be repeated in each game by pressing a button labeled with a question mark.

2.3. Implementation of the selected games

The first implemented task is ‘Reading with Arcs’ in the analogue reading intervention. Children are instructed to mark the syllables of words with arcs and thereby read the words syllable-by-syllable. In the analogous intervention, syllables are marked with arcs while reading a word aloud with the whole class, as well as silently during individual work. The aim is to make syllables the salient unit of
Towards a digital syllable-based reading intervention
Schaper et al.

processing instead of reading words letter by letter (Müller, Richter, & Karageorgos, 2020).

The game ‘Reading with Arcs’ from the analogue intervention was adapted and called ‘Flower Meadow’ and is associated with the flower button in the selection screen. Similar to the analogue version of the game, users have to hyphenate words into syllables. Separation of the words is represented by the flight path of a bee. Every letter of the respective word is placed on a separate flower (Figure 1). Users separate the syllables by dragging the icon of a bee to the final letter of the current syllable within the word. The number of syllables per word varies according to the level difficulty of the game. Visual and audio feedback is used to indicate a right or wrong answer. Once users correctly hyphenate a word, their progress is acknowledged and they proceed with the next word in the level. Difficulty is adapted by inclusion of longer words in higher levels. Once the progress bar in the level is filled, the user is informed about passing the game in the respective difficulty and allowed to return to the difficulty selection screen.

The second selected game is ‘Vowel and Consonant’, which serves primarily to teach the composition of a syllable and the distinction between vowels and consonants. The original version of the learning exercise describes each syllable as a boat, which has exactly one captain but several sailors. The vowel, or diphthong, is the captain of the boat, the consonants are the sailors that the children must recognize. In German each syllable contains a vowel or diphthong as its nucleus, therefore identifying the vowel or diphthong can help children to distinguish the syllables within a word. This general rule is introduced in the third session of the analogue intervention and is trained in every session afterwards. Recognizing individual syllables and their structure is a fundamental requirement for the differentiation of syllables and thus the reading of whole words.

In the digital version ‘Vowel and Consonant’ was adapted and renamed as ‘Sailor Game’ and is associated with the lighthouse button in the selection screen. The digital version visualizes the explanation used in the reading intervention of each syllable as a boat with a captain by dividing a word and placing its syllables in separate boats, initially representing each letter as a sailor (Figure 2). Users have to recognize the vowels, vowel mutations or diphthongs as the captain of each syllable of the depicted word. The goal for each level is to find all captains by touching the vowels and diphthongs. Positive feedback is shown by highlighting the correctly selected figures with a green border and changing the sailor into a captain, negative feedback by a red shading, and a shake animation. Furthermore, the facial expression of the selected sailor changes accordingly as well as the auditory feedback. The difficulty of each level varies due to the inclusion of monosyllabic words in difficulty level one, progressively to three-syllable words in difficulty level three. Once the progress bar in the level is filled, the user is informed about passing the game in the respective difficulty and allowed to return to the difficulty selection screen.

The third game ‘Syllable Salad’ occurs first in the sixth session of the reading intervention. In the original version of the intervention, the child receives an envelope, which contains cards with several separate syllables. The task is to bring these syllables in the correct order to merge them into words. In the digital version, ‘Syllable Salad’ is represented as a school building on the overview map. Users rearrange individual syllables to form a complete word. To this end, the syllables of a word are displayed randomly scattered on a blackboard (Figure 3). The task is to place the syllables in the right order by using a drag and drop mechanic. Positive and negative feedback is communicated via sound output. The individual levels vary in their degree of difficulty in terms of the number of syllables. In difficulty level one, two-syllable words are presented, while difficulty level two consists of three-syllable and difficulty level three of four-syllable words. Once the progress bar in the level is filled, the user is informed about passing the game.
in the respective difficulty and allowed to return to the difficulty selection screen.

![Image](image_url)

**Figure 3: The game ‘Syllable Salad’ during gameplay.** *(The letters show the word Giraffen / giraffes).*

### 2.4. Expert review

In order to examine whether the four requirements for the app (efficiency, effectiveness, learnability and child-friendly design to assure user satisfaction and reduce cognitive load) have been met by our implementation, and prior to conducting a study with second-graders, the prototype was evaluated by experts. Three researchers (one professor, 2 postdocs) at the chair of educational psychology separately used and commented on the app. In individual interviews, the experts first watched and commented the instruction videos. Afterwards, they were invited to use the app without time limit and to freely navigate within the app. During interaction with the app, the experts gave oral feedback, which was recorded. After they finished operating with the app, additional half-structured interviews were conducted. The questions focused on possible difficulties that could arise during the children's app usage, issues the pupils could need assistance with, and the adequacy of the implemented feedback. They were also asked to provide feedback for the procedure of the study based on their expertise, and to comment on the interview guideline for half-structured interviews, planned to be conducted with the children in the subsequent study.

Regarding the instruction videos, the experts mainly reported potential problems with the wording of the conveyed information and the feedback. The experts provided recommendations for the graphical user interface, especially for the main menu's design, the font size, as well as the font type within the games. They further suggested to include an instruction video to explain the general handling of the app. Finally, the usability of the game ‘Syllable Salad’ was criticized. All expert feedback was taken into account and the app was adapted accordingly.

After implementation of the expert feedback, we presented the app to the class teacher of the pupils participating in the study, as well as an additional primary school teacher. Both explicitly confirmed the app's and interview's suitability for the evaluation.

### 3. USER TESTING

We conducted a study with a school class of second-graders who interacted with the app in a controlled setting and took part in a subsequent interview. The main aims were to test if the central aspects of the analogue syllable-based reading intervention can be successfully transferred into a digital version (RQ1), whether second-graders rate the app as motivating (RQ2) and whether second-graders enjoy using the app (RQ3). For future adaptations, we were interested in the performance of the children using the app and therefore measured errors and time spent on specific tasks.

#### 3.1. Participants

The study was conducted in a primary school in Germany. In total, 17 children with written consent by their legal guardians used the app. The age ranged from seven to nine years ($M = 7.65$, $SD = 0.61$). Twelve girls with a mean age of 7.75 years ($SD = 0.62$) and five boys with a mean age of 7.40 years ($SD = 0.55$) participated in the survey. According to the class teacher, the reading abilities of the participating children were heterogeneous.

#### 3.2. Design and procedure

The study was separated into two parts for each participant, with two investigators assigned to one child. The sessions were held individually and were set to not exceed 40 minutes.

First, each pupil was invited to interact with the app presented on a tablet, playing each game for approximately five minutes, using the same sequence of the three games, ‘Flower Meadow’, ‘Sailor Game’ and ending with ‘Syllable Salad’. Afterwards, each child could freely use the app for up to five more minutes. The screen capture software ‘DU Recorder’ (DU-Apps-Studio, 2018) was used to record the complete interaction with the app. Questions asked during the interaction and the children's behaviour were recorded manually in writing, including instances in which the child needed assistance.

Second, a half-structured qualitative interview was conducted based on the interview guidelines discussed with the experts. One of the investigators asked questions and interacted with the participant while the other researcher noted answers and behaviour. The interview consisted of 17 open-text questions with child-friendly wording, which concerned demographic information, including age and gender, experience with tablet usage, joy of use, perceived learning outcome, basic understanding of the app’s handling, the app’s design and the design of the individual games. The pupils were asked which of
Towards a digital syllable-based reading intervention
Schaper et al.

the games they perceived to be hardest and easiest respectively, which game they liked best and which game they wanted to play again in the future. There were also open questions for positive and negative feedback on each individual game. Finally, the pupils were asked for general feedback and were invited to give additional suggestions.

4. RESULTS

The interviews were evaluated using the data and text analysis software MAXQDA 2018 (VERBI Software, 2018). Statistical tests were conducted using SPSS Version 25.

4.1. Screen-capture data

Due to technical problems, only 16 screen-capture videos could be analysed. Each video was evaluated independently by two investigators. The following criteria were extracted from each screen capture: (1) number of accomplished levels within five minutes, (2) time per level, (3) number of content-related errors, (4) number of interface-related errors and (5) choice of games during free interaction. Errors were separated into two categories to differentiate between problems with the interaction (interface-related), such as trying to interact with a non-interactive object, and problems with the reading task (content-related), such as selecting a wrong answer (Table 1).

Table 1: Time per level, content-related and interface-related errors for all three games and their respective levels.

<table>
<thead>
<tr>
<th>Game</th>
<th>Level</th>
<th>Time per level Mean (SD) Min - Max</th>
<th>Number of content-related errors Mean (SD) Min - Max</th>
<th>Number of interface-related errors Mean (SD) Min - Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flower Meadow</td>
<td>Level 1 (n=16)</td>
<td>1.32 (0.38) 0.53-2.59</td>
<td>2.88 (0.38) 0-5</td>
<td>2.69 (0.07) 0-11</td>
</tr>
<tr>
<td></td>
<td>Level 2 (n=15)</td>
<td>1.29 (0.40) 0.50-3.39</td>
<td>4.87 (0.60) 0-5</td>
<td>1.60 (0.20) 0-3</td>
</tr>
<tr>
<td></td>
<td>Level 3 (n=8)</td>
<td>1.26 (0.18) 0.59-1.58</td>
<td>5.11 (0.45) 0-4</td>
<td>0.56 (0.19) 0-2</td>
</tr>
<tr>
<td>Sailor Game</td>
<td>Level 1 (n=16)</td>
<td>0.57 (0.40) 0.16-2.40</td>
<td>2.69 (0.24) 0-3</td>
<td>0 0</td>
</tr>
<tr>
<td></td>
<td>Level 2 (n=16)</td>
<td>0.40 (0.16) 0.18-1.19</td>
<td>1.56 (0.09) 0-3</td>
<td>0 0</td>
</tr>
<tr>
<td></td>
<td>Level 3 (n=15)</td>
<td>0.46 (0.12) 0.25-1.18</td>
<td>1.40 (0.08) 0-3</td>
<td>0 0</td>
</tr>
<tr>
<td>Syllable Salad</td>
<td>Level 1 (n=16)</td>
<td>0.44 (0.07) 0.35-1.03</td>
<td>0.69 (0.13) 0-2</td>
<td>1.06 (0.15) 0-4</td>
</tr>
<tr>
<td></td>
<td>Level 2 (n=16)</td>
<td>1.23 (0.28) 0.42-2.29</td>
<td>3.81 (0.55) 0-7</td>
<td>0.69 (0.07) 0-2</td>
</tr>
<tr>
<td></td>
<td>Level 3 (n=15)</td>
<td>2.04 (0.25) 1.25-2.54</td>
<td>8.53 (0.93) 0-8</td>
<td>1.43 (0.19) 0-4</td>
</tr>
</tbody>
</table>

Note: Standard deviations are in parentheses.

Interface-related errors should be as low as possible while content-related errors are based on the ability of the learner and are not necessarily zero when confronted with more difficult tasks. Each child completed all levels of 'Syllable Salad' as well as all levels of 'Sailor Game', except for one participant who did not finish the third level before the five minutes had elapsed. In ‘Flower Meadow’, each participant finished the first level, but only nine children completed the third level within the five minutes.

The mean time spent on each level of ‘Flower Meadow’ was 01:29 minutes, for ‘Sailor Game’ it was 00:48 minutes per level and for ‘Syllable Salad’ the mean duration was 01:23 minutes per level. Only for the game ‘Syllable Salad’, the mean time increased with difficulty of the levels. In ‘Flower Meadow’ the average time per level was nearly the same for all three levels while in the sailor game the first level took longest (Figure 4).

Figure 4: Average time required for each completed level within each game. Error bars represent SEs.

To assess content und interface-related errors we calculated the mean number of errors, only for completed levels, within each game, resulting in a
decrease in participants in higher levels (see Figure 5).

Descriptively, the ‘Sailor Game’ had the lowest number of content-related errors overall, followed by ‘Flower Meadow’ and finally ‘Syllable Salad’ with the highest number of content errors. It is noteworthy that ‘Sailor Game’ had the highest error rate in the first level whereas in ‘Flower Meadow’ and ‘Syllable Salad’ the content-related errors increased with the difficulty of the levels. Concerning the interface-related errors, ‘Sailor Game’ had no interface-related errors at all. ‘Syllable Salad’ had fewer interface-related errors compared to ‘Flower Meadow’.

When being allowed to play with the app freely, two of the children were not interested in more interaction with the app, and for three participants data was not recorded during their interaction. The remaining pupils (n = 13) played one or multiple games during their free interaction. ‘Flower Meadow’ was played by ten children, ‘Sailor Game’ by eight children, and ‘Syllable Salad’ was chosen by six children.

![Figure 5: Mean content errors and mean interface-related errors per child and level. Error bars represent SEs.](image)

4.2. Interview data

All interview data was separated into segments, which were assigned a category by a group of three raters based on majority. The same three raters assigned categories for all interviews.

Regarding the app in general, all children reported having fun while using the app and separately claimed that they liked the appearance of the app. Concerning the overall interaction, 13 participants stated that the handling of the app was clear and comprehensible to them.

When asked about the enjoyment of each game separately, all children affirmed that they enjoyed playing the ‘Sailor Game’. The pupils gave different reasons why they enjoyed the game, such as the game being easy to solve, easy to use, funny, that they liked the sailor theme and the visual feedback. For ‘Flower Meadow’, 15 children stated that they enjoyed the game, with the main reason of having fun playing the game. For ‘Syllable Salad’ also 15 children affirmed that they enjoyed playing the game and stated having fun and handling the game easily. Additionally, four children also mentioned the learning effect and the challenge the game posed.

Regarding the perceived difficulty, six children chose ‘Syllable Salad’ as the most difficult game, while ‘Flower Meadow’ and ‘Sailor Game’ were mentioned four times each. As the easiest game, ‘Sailor Game’ was chosen seven times, ‘Flower Meadow’ six times and ‘Syllable Salad’ three times.

When the participants had to decide on their favourite game within the app, six of the children chose ‘Flower Meadow’, followed by five children selecting ‘Syllable Salad’ and four of the children choosing ‘Sailor Game’. Two children had no favourite game. The children stated that their choice was based on the easy handling, easy solvability, fun and overall appearance.

When asked whether they would like to use the app again, all children agreed. However, ten children were only interested to replay certain games: ‘Syllable Salad’ as well as ‘Sailor Game’ were mentioned seven times, ‘Flower Meadow’ was named three times.

Interestingly, 14 children stated that they have learned something while playing with the app. Improving in splitting syllables was mentioned by eleven children, four stated that they improved their understanding of words and vowels, and two said to have improved in reading.

When asked for general feedback on the app, four of the children emphasized the graphics, three mentioned having fun and two commented on the perceived learning. Additionally, three pupils mentioned the simple handling and two enjoyed the challenge of the tasks. Regarding negative aspects, one of the children did not like the type of visual feedback and one noted that the difficulty might not be suited based on individual abilities.

Last, the pupils were asked for suggestions on how to improve the app. In general, improvement of graphics was mentioned four times. Additionally three of the children asked to implement more games and levels. Most suggestions for improvement occurred on ‘Syllable Salad’. Individual suggestions included to use funny words, increase difficulty and use different graphics. In total, having fun while using the app was mentioned 31 times across all interviews. The easy interaction was named 15 times and the perceived knowledge gain ten times.

5. DISCUSSION

Concerning RQ1, the evaluation by psychological experts, teachers and the study demonstrated that
the selected games of the analogue training can be digitalized. The pupils were overall able to independently work with the app based on the integrated instructions. However, the results have to be considered in detail. The quantitative data demonstrated substantial variability within the participating children in the abilities needed for the content of the app. There was also high variability in the interface-related errors for ‘Flower Meadow’ and ‘Syllable Salad’. However, content-related errors were more numerous than interface-related errors in all three games. Only in the ‘Sailor Game’ no interface-related errors occurred, indicating the need to improve the interaction within both other games. For RQ2 and RQ3 we suggest that the results, especially the interviews, might have been impacted by the novelty of the games itself, especially relative to day-to-day school activities.

The interaction with the app itself should be unaffected in terms of content-related and interface-related errors and we still consider the information gained in the interviews as valuable insights into the target group. The overall reception of the app and the three implemented games was very positive. Given the fact that the games are clearly focused on learning, the interest of all participating children to use the app again is an encouraging sign, supporting the notion that the app is perceived as motivating, even though it should be treated with caution. Furthermore, the heterogeneous choice as favourite game can be considered as an indicator that the games are comparably enjoyable, which is supported by how often the children reported having fun, which supports RQ3.

5.1. Lessons learned

However, the performance of the children within the app in terms of time per level, content-related and interface-related errors indicates room for improvements. The time per level, as well as content-related errors increased in higher levels in ‘Syllable Salad’, with only few interface-related errors in each level. This suggests that the planned increase in difficulty across levels worked and shows that there appear to be no significant problems in using the game.

For ‘Sailor Game’ the mean time per level is similar across levels and the low mean number of content-related errors is even decreasing in higher levels. There were no interface-related errors at all for this game. ‘Sailor Game’ can therefore be considered as a relatively easy game in terms of content, which is also easy to use, as supported by the interview data.

Concerning ‘Flower Meadow’, the mean time per level is also constant across levels and content-related errors increased slightly in higher levels. Both measures can be considered as an indicator that the planned increase in difficulty across levels was noticeable. However, the number of interface-related errors in ‘Flower Meadow’ is descriptively higher compared to both other games, especially in the first level. We suggest that, even though the game appears to be not very difficult in terms of content, the current mode of interaction is not ideal, which results in increased interface-related errors. The interaction should be improved, especially because the games were selected to be suited as introduction for the app and therefore should be easy to interact with. Furthermore, these interface-related errors might be perceived as content-related errors by the children, potentially resulting in confusion about the correct solution.

The difficulty of ‘Syllable Salad’ and ‘Flower Meadow’ measured in errors matches the subjectively perceived difficulty of the games as indicated by the choices for easiest and most difficult game. Interestingly the larger amount of input-related errors in ‘Flower Meadow’ did not appear to have influenced this assessment. Furthermore, considering the choice of favourite game, ‘Flower Meadow’ was chosen most often. The input-errors might therefore not be regarded as a critical problem for the children, this is most likely the case because errors were not costly and the correct input could be executed without delay.

When the children could freely interact with the app, half of the children chose the levels of the games they did not manage to finish before. This behaviour can be interpreted as an indication that the children were motivated to finish all levels. The implemented aspects of gamified elements in the learning environment such as the progress bars within each level and the increasing difficulty across levels might have incentivised the children to rise up to these challenges. In summary, the degree of difficulty appeared appropriate and motivating for the children and the overall interaction with the app worked well.

Considering the initial research aim of adapting an analogue reading intervention to a mobile app, several aspects can be noted. Foremost, based on the initial prototype and the reactions within the study, it can be assumed that central elements of the intervention can be implemented into an app (RQ1). Furthermore, the results of the interviews illustrate the children’s motivation to use the app again, (RQ2). The interviews also show an overall enjoyment of the children using the app (RQ3). The results should be considered in light of a relatively small group of participants in a situation that is not equivalent to the planned context of use. We still regard these preliminary indications for future development as highly promising.

The combined approach of capturing and analysing the interactions with the games and the use of the interviews allowed to identify problems as well as potentials. This allows us to address the problem of input-errors in a more differentiated way, seeing that even though minimizing these errors is an
obvious task in development, the consideration of a design, which minimizes their impact is also worthwhile for subsequent versions.

5.2. Design implications

Based on the results of the interviews and analysis of the screen-capture data we derived several concrete improvements for the app that might also serve as general design implications for app development in this area. Regarding the app’s user interface it is noticeable, that the use of text requires a font that is easy to read and therefore more suitable for beginning readers. Furthermore, the function of the buttons was not always clear. Although the buttons were already designed to be noticeable in size and contrast, it can be generally noted that app developers should focus more on the layout of critical user interface elements, such as buttons. This includes their self-explanatory presentation in the app and intuitive use. This is in line with the need for help by some children throughout the study. Therefore, the tutorials and help functions have to be extended to improve understanding. These observations indicate that especially children might profit from a more interactive way of explaining the core app functions, including active participation and exercises.

Furthermore, our results highlight the need of a stepwise and differentiated approach for transferring an analogue concept into a digital application. The selected requirements (1) efficiency in terms of adequate and accurate feedback, (2) effectiveness with regard to the comprehensibility of content and navigation elements, (3) the self-descriptiveness of the individual games and the app overall (learnability) and (4) child-friendly design to assure user satisfaction and reduce cognitive load can only be addressed by an iterative process including experts and the target group of children. Expert feedback alone, which is often used in form of heuristic evaluations has clearly not been sufficient to finalize the app as demonstrated by the study. However, the study with the second-graders by itself might have been similarly problematic because central problems such as the approach to the instructions and potential difficulties with the initial font had to be identified to actually allow for meaningful interaction with the app.

6. CONCLUSION

For this contribution, we used a human-centred design approach to develop a prototype of a mobile learning app based on an analogue syllable-based reading intervention (Müller, Richter, & Otterbein-Gutsche, 2020). Such a digital version could help to address the need for a more accessible reading intervention for children with deficient reading abilities.

The preliminary prototype was subject to an evaluation by experts and was subsequently revised. A study was conducted with a class of second-graders, who played with the app and were interviewed about the interaction. In general, the participating children enjoyed using the app and were motivated for continued use. The study demonstrated that core aspects of the analogue reading intervention can be adapted into a digital version. Results of the study generated valuable insights to improve the current state of the prototype. The feedback from the target age group can help guide future development of the app, e.g., on how to improve the introduction tutorials, therefore ensuring the quality of the developed app (Papadakis et al., 2018). These insights also provide implications regarding the development and design of educational apps for children.

Engaging a human-centred design approach allowed to incorporate the perspective of theoretical experts (i.e., pedagogical psychologists), practical experts (i.e., primary school teachers) and users (i.e., second-graders). This enabled us to iteratively improve the prototype, establish a solid foundation for subsequent development of the prototype into a complete app, and to present this development approach as insight for similar endeavours.

The participating children also asked for extension of the levels and the app itself, which is less a needed improvement and more an indication that there is demand for the adaption of the reading intervention by Müller et al. (2018).

In the long term, the learning outcome of the app needs to be examined with pre and post tests for children with different reading abilities in a large sample. Because the underlying intervention has been shown to be effective (Müller, Richter, & Karageorgos, 2020; Müller et al., 2017) in a long term study, the effect of such a short interaction as in the study is only the first step in a comprehensive evaluation of the digital adaption. The goals during development were focused on our research questions: the proof of concept, motivation in the target group and interest in use of the app. Our goal is therefore to scientifically establish a usable and well thought out app, which adequately represents and meaningfully expands the analogue intervention, and also test this improved version in terms of learning success on a large scale (Haßler et al., 2016).

7. ACKNOWLEDGEMENTS

This work was funded by the German Federal Ministry of Education and Research (BMBF) under the grand agreement MobiLe (03VP07080).
8. REFERENCES


