SpARklingPaper: Enhancing Common Pen- And Paper-Based Handwriting Training for Children by Digitally Augmenting Papers Using a Tablet Screen

TOBIAS DREY, Institute of Media Informatics, Ulm University, Germany
JESSICA JANEK, Institute of Media Informatics, Ulm University, Germany
JOSEF LANG, Institute of Media Informatics, Ulm University, Germany
DIETMAR PUSCHMANN, Institute of Media Informatics, Ulm University, Germany
MICHAEL RIETZLER, Institute of Media Informatics, Ulm University, Germany
ENRICO RUKZIO, Institute of Media Informatics, Ulm University, Germany

Educational apps support learning, but handwriting training is still based on analog pen- and paper. However, training handwriting with apps can negatively affect graphomotor handwriting skills due to the different haptic feedback of the tablet, stylus, or finger compared to pen and paper. With SpARklingPaper, we are the first to combine the genuine haptic feedback of analog pen and paper with the digital support of apps. Our artifact contribution enables children to write with any pen on a standard paper placed on a tablet's screen, augmenting the paper from below, showing animated letters and individual feedback. We conducted two online surveys with overall 29 parents and teachers of elementary school pupils and a user study with 13 children and 13 parents for evaluation. Our results show the importance of the genuine analog haptic feedback combined with the augmentation of SpARklingPaper. It was rated superior compared to our stylus baseline condition regarding pen-handling, writing training-success, motivation, and overall impression. SpARklingPaper can be a blueprint for high-fidelity haptic feedback handwriting training systems.

CCS Concepts: • **Human-centered computing** → *Tablet computers*; **Mixed / augmented reality**; *User interface design*; • **Social and professional topics** → **K-12 education**; • **Applied computing** → **Education**.

Additional Key Words and Phrases: artifact; augmented reality; mobile devices; tablet; literacy training; handwriting training; children; education

ACM Reference Format:
Tobias Drey, Jessica Janek, Josef Lang, Dietmar Puschmann, Michael Rietzler, and Enrico Rukzio. 2022. SpARklingPaper: Enhancing Common Pen- And Paper-Based Handwriting Training for Children by Digitally Augmenting Papers Using a Tablet Screen. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 3, Article 113 (September 2022), 29 pages. https://doi.org/10.1145/3550337

1 INTRODUCTION

How children learn to write in school has not changed in the last years, although digital tools such as tablets and smartphones are everywhere in our daily lives. Pen and paper are still the tools used and very important

Authors' addresses: Tobias Drey, Institute of Media Informatics, Ulm University, James-Franck-Ring, Ulm, 89081, Germany, tobias.drey@uni-ulm.de; Jessica Janek, Institute of Media Informatics, Ulm University, James-Franck-Ring, Ulm, 89081, Germany, jessica.janek@uni-ulm.de; Josef Lang, Institute of Media Informatics, Ulm University, James-Franck-Ring, Ulm, 89081, Germany, josef.lang@uni-ulm.de; Dietmar Puschmann, Institute of Media Informatics, Ulm University, James-Franck-Ring, Ulm, 89081, Germany, dietmar.puschmann@uni-ulm.de; Michael Rietzler, Institute of Media Informatics, Ulm University, James-Franck-Ring, Ulm, 89081, Germany, michael.rietzler@uni-ulm.de; Enrico Rukzio, Institute of Media Informatics, Ulm University, James-Franck-Ring, Ulm, 89081, Germany, enrico.rukzio@uni-ulm.de.

2474-9567/2022/9-ART113
https://doi.org/10.1145/3550337

Fig. 1. We present SpARklingPaper, a system for handwriting training for elementary school children, based on a combination of a pen, a paper, and a tablet. First, the paper is placed on the tablet, fixed with magnetic stripes, and augmented from below by the tablet's screen shining through it (a). Next, letters are introduced with a big capital letter and a related animal (parrot for 'P'; "Papagei" in German). Then, the children can practice the letter by following the strokes (b). Finally, we provide individual and immediate feedback shown in green and red if a letter was written correctly (c,d).

for developing fine graphomotor handwriting skills [2]. Nevertheless, as still, 12% to 30% of children fail to acquire the motor skills for handwriting [15], which results in >30% of the girls and >50% of the boys having problems with fluid and legible handwriting [51], it is an important question how to improve learning how to write. This learning deficit is currently increasing dramatically due to COVID-19 pandemic-related school closures, resulting in an estimated loss of 20% of a school year and only a little or close to zero learning progress during homeschooling [23]. This learning loss also negatively impacts long-term learning when being back at school [39].

Using digital devices can support learning [16], as, e.g., tablets can improve children's motivation and concentration for literacy training and language skills [11, 26, 52, 57, 79]. Apps can provide individual and instant feedback, which a teacher cannot do while teaching class [8]. This can further increase self-paced learning, foster the flow state, and prevent children from being frustrated or stuck [11, 17, 19, 20]. Besides tablet apps, visual support based on augmented reality (AR) is another approach to enriching learning [34, 35, 56, 59, 61]. Learning applications provide further the possibility to educate children at home, e.g., during COVID-19 pandemic, or to compensate for resulting knowledge gaps.

Despite all these advantages, graphomotor handwriting skills can be negatively affected when writing is not trained traditionally with pen and paper [2]. This can be explained by the lower friction of a tablet screen when writing on it with a stylus or a finger [2, 12, 13, 27, 30, 71]. A further issue is the transfer of skills as it is not trivial to practice with a stylus and a tablet and then write with the everyday standard pen and paper [3, 42, 66]. Therefore, using pen and paper and augmenting it seems to be a promising approach. This allows building systems based on traditional tools, which have no skill transfer issue and provide the same educational benefits as digital devices and apps. Projectors were used for this previously for literacy/handwriting training [34, 35, 56, 61], but the child's hand and pen would occlude the augmentation. Also, head-mounted displays (HMDs) were used for literacy/handwriting training of adults [63, 64] but are not approved for preschool and elementary school children [50, 55].

We present SpARklingPaper, a system combining a pen, a paper, and a tablet for augmented handwriting training, which combines the graphomotor benefits of the genuine haptic feedback of pen and paper with the educational benefits of interactive and individual digital learning on tablets. It solves the occlusion issues of projectors and is unlike HMDs usable by children. SpARklingPaper is an artifact contribution according to Wobbrock and Kientz [75].

Our design goals, which were created together with our advising experts, one teacher and one mother of elementary school and kindergarten children, were,

(1) creating a system where children can use the pen of their choice to write with digital support on standard paper,
(2) creating a system based on common and cheap hardware that is affordable for everyone,
(3) creating a mobile system usable with multiple simultaneous instances in class and as a single instance for homework or homeschooling,
(4) and providing the possibility to non-developers to create their own learning content for our prototype.

In SpARklingPaper, a standard paper is placed on the tablet's screen, and the children can write with a pen of their choice on it. The tablet can augment this paper by shining through it (see Figure 1a). Letters are introduced with an animation showing the order of the strokes. Children can then follow these strokes and practice the letters (see Figure 1b). Helping lines for novice writers are further shown. Our system also provides individual and immediate feedback and shows the children with green and red areas if they have written a letter correctly (see Figure 1c/d). This way, children can fix their mistakes very fast and do not have to wait for feedback from a teacher or parent. We use a connected smartphone on a stand that could capture the written letters to evaluate the writing as the touchscreen cannot recognize the pen-based input (see Figure 2). Besides the concept and the prototype, we present the user-centered design process accompanied by two external advisers (an elementary school teacher and a mother of elementary school and kindergarten children) and example learning content for the prototype based on the current curriculum of first- and second-grade elementary school in the country of residence of the authors.

To evaluate our prototypical implementation of our SpARklingPaper concept against a state-of-the-art stylus- and tablet-based baseline condition, we created a research question (RQ) considering the previously discussed important influences for writing training: *"Does the combination of a pen, a paper, and a tablet for writing training, as SpARklingPaper provides it, impact pen-handling, assumed writing training-success, usability, motivation, feedback understandably, perceived functionality, and the overall system impression, compared to a state-of-the-art stylus- and tablet-based training?"*

To evaluate our system and answer our RQ, we conducted a twofold evaluation based on an initial online survey (N=29: 25 parents, 4 teachers) followed by an in-person user study (N=26: 13 children, 13 parents). We created two online surveys for teachers (44 questions) and parents (63 questions) of a partnering elementary school and introduced our SpARklingPaper prototype with two videos, and asked general questions regarding children's writing skills, state-of-the-art writing education at school, homework, and children's motivation to learn writing. This survey was followed by an in-person user study comparing SpARklingPaper against a state-of-the-art stylus-based condition, as described in our RQ, based on questionnaires and semi-structured interviews for parents and children.

Our survey findings show that there are several needs to improve handwriting training. First, the writing skills of children are at the beginning of first-grade relatively low, and they have to invest a lot of time and effort during first- and second-grade conducting swing exercises and learning the strokes of block and cursive letters as well as acquiring a proper hand posture for holding the pen, which shows the high potential of SpARklingPaper at school. Second, children's motivation during homework drops fast. Therefore, parents stated that they are grateful for a system that supports and motivates their children. Our study results show that SpARklingPaper can be such a system by approving our design concept of combining pen and paper's genuine analog haptic feedback for handwriting training with AR. Both children and parents rated it to provide better pen-handling based on the own common pencil than the stylus baseline, which they assume would result in superior writing training-success as well as higher motivation. This resulted in an overall better impression than the stylus baseline, and SpARklingPaper would be preferred over classic analog pen- and paper-based writing as well. They further

rated our individual augmented feedback with green and red areas as advantageous for class and homework. Despite the prototype of SpARklingPaper needing more hardware and a more complex setup than the stylus baseline, the usability was not dramatically affected, and parents and children would accept this for an improved learning experience. These findings show that our SpARklingPaper concept and our prototype are a blueprint for future high-fidelity haptic feedback handwriting systems.

The main contributions of our work are:

(1) The design of an augmented handwriting system concept called SpARklingPaper, based on the novel combination of a pen, a paper, and a tablet in one integrated system considering four design goals as artifact contribution, based on an iterative user-centered design (UCD) process with a teacher, a mother, and their children.
(2) The implementation and presentation of our tablet- and smartphone-based prototype, including the presentation of an example learning unit.
(3) The conduction and evaluation of two user surveys (N=29) with >100 questions targeting the current situation of handwriting training at school as well as our SpARklingPaper concept and its potential impact on class and homework.
(4) The conduction and evaluation of an in-person user study (N=26) with 13 children and 13 parents to compare SpARklingPaper against a state-of-the-art stylus- and tablet-based baseline condition regarding pen-handling, assumed writing training-success, usability, motivation, feedback understandably, perceived functionality, and the overall system impression.

2 RELATED WORK

Our work is based on previous work regarding tablet-based literacy/handwriting training and AR supported writing. We will discuss the benefits and weaknesses of these approaches and how they influenced our work with its goal to combine their benefits.

2.1 Tablet-based Literacy/Handwriting Training

Digitally supported literacy training is beneficial for pupils. Gamified learning can improve the motivation and engagement of learners [11, 52], as well as automatic feedback, which could support or replace a teacher [8]. These benefits can also be transferred to calligraphy training, which is highly related to letter handwriting training [49, 76, 77]. Due to these benefits, many commercial apps exist in the app stores of Apple and Google that promise to help to learn writing [21, 32, 65]. Several of them can be categorized in apps that teach writing by showing huge letters, which should then be written with a stylus or a finger [7, 45, 62, 73]. Other apps also exist, which focus on reading [1, 24], grammar [29, 33], and spelling [29, 33] training, but we will not focus on them as this work focuses on teaching the mechanical skill of writing by hand.

Typical literacy training research investigates different digital systems against each other or a baseline condition with pen and paper. Patchan and Puranik [58] compared three ways of learning writing: pencil and paper, tablet and finger, and tablet and stylus. The preschool children were training eight letters for eight weeks. In a post-test, conducted with pencil and paper, they had the result that the amount of correctly written letters is similar when children write with a stylus on a tablet or with a pencil on paper. However, the children using the tablet performed less than half the amount of trials to have a similar post-test result. The best results for correctly written letters had the children who wrote with a finger on a tablet. Their findings indicate the benefits of using tablets in literacy training for children. Similar positive effects were found by Bonneton-Botté et al. [8], who investigated the effectiveness of a stylus-based tablet app for preschool children handwriting training during a twelve-week period. They analyzed the effect of automatic feedback delivered at the end of each trial to inform learners of

their results. Their study showed that a stylus-based tablet app with automatic feedback can support the learning of handwriting compared to traditional pencil- and paper-based education, assessed with a pencil and paper test.

Nevertheless, previous works also state negative results for digitally supported literacy training. Mayer et al. [53] and Kiefer et al. [41] investigated the influence of different writing tools on kindergarten children's reading and writing performance. Over seven weeks, reading and writing 16 letters and twelve words were trained with either pencil and paper, tablet and stylus, or a virtual tablet keyboard. The reading and writing performance was assessed before training, after training, and four to five weeks after training with standardized tests. The test was performed with the training system, not with pen and paper. They conclude that writing with a stylus on a touchscreen is the least favorable approach, as the demand for motor control increases due to the low friction of the touchscreen glass compared to paper. Similar results were found by Gerth et al. [27], who compared the influence of the writing surface on handwriting performance for non-writers, beginners, and experienced writers. They compared pen and paper with stylus and tablet and showed that all groups were affected by the lower friction of the touchscreen. They conclude that it is not recommended to use tablets in school for writing as the lower friction of a tablet screen needs additional control of handwriting movements, which is an additional challenge for learners. Guilbert et al. [30] also observed that the friction of the touchscreen could be a destabilizing factor of handwriting on a tablet screen for children and advised to increase the friction of the screen, which our work proposes a solution for. Further negative effects of the lower friction writing surfaces are slower writing speed [12], using more pressure during writing [2, 71], and reduced legibility [2]. These works show the benefit of our approach, as we combine pen and paper and a tablet and therefore their benefits by simultaneously excluding their drawbacks.

Another potential issue of practicing with a stylus, finger, or keyboard on a tablet is how this training transfers to the everyday standard pen and paper. This is linked to theories of how well one skill transfers to another area [3, 42, 66] and was the reason why some previous studies did not investigate a skill transfer and tested their participants with the training system (see Mayer et al. [53]), or the skill transfer could be the reason why the stylus- and tablet-based system did not perform better than the baseline pen and paper (see Patchan and Puranik [58]). As SpARklingPaper proposes to train with pen and paper, we do not have a skill transfer issue.

2.2 Augmented Reality Supported Writing

Previous works have created AR approaches to augment real objects to use hardware other than a tablet or a stylus for literacy or calligraphy training. Shichinohe et al. [61] created a system based on a projector, a partly transparent glass, and a brush for calligraphy training. An approach used multiple times in literature [34, 35, 56]. The calligraphy was augmented with the projector, which was projecting on the glass put on a table. As the glass is quite similar to a touchscreen glass, the friction was different than on paper, but this was not the focus of their work, as they were also supervising the participants with cameras to give them feedback about their body posture, which influences the quality of the calligraphy.

To support augmented writing on paper, Winkler et al. [74] created a combination of a tablet augmenting a paper placed on a desk with a compact projector as a prototype to support medical doctors in digitizing their notes. They used the Anoto pen and paper [4] to track the pen, and the system did not write with ink but projects the writing. The AR Lamp by Kim et al. [43] also uses a projector to augment paper placed on a desk and adds dynamic information. The pen's writing is recognized by a camera combined with an Anoto pen and paper. Their prototype was very generalizable and usable, e.g., for music theory education, and focused as well on finger gesture-based interaction. Similar systems based on projectors projecting on arbitrary surfaces were proposed by Xiao et al. [78], Wellner [72], Hardy [31], Linder and Maes [47], Junuzovic et al. [38], Kane et al. [40], and Underkoffler et al. [68]. Some of these systems used finger and gesture input, and not all supported a stylus or a pen. While these systems seem similar to our approach, they had other goals and were primarily technical

demonstrators not focusing on literacy/handwriting training or the benefits of the haptic feedback of pen and paper for handwriting training. Using a projector has further the drawback that the own hand, arm, and body can occlude the projected surface and break the system. Using a tablet to augment the paper also allows a more compact and mobile setup than using a projector.

An HMD can solve the occlusion issue of the own body, too. Soontornvorn et al. [63, 64] created such a system for calligraphy learning, and HoloDoc [46] is an example of an interactive desk based on an HMD. Whereas HMDs are quite powerful, our system is intended for handwriting training of preschool and elementary school children, which are between 4-8 years old, and the common devices such as HoloLens 1 and 2 as well as Magic Leap are only approved for children older than 13-14 due to their interpupillary distance [50, 55] and therefore not usable by the target group.

Other works exist, which focus on pen-based input in mixed reality (MR) [5, 6, 18, 48, 60]. These systems used or built specialized pens to improve input capabilities in MR, but their focus was neither on literacy/handwriting training nor on providing the haptic feeling of writing with an ordinary pen on paper.

3 THE SPARKLINGPAPER SYSTEM

With SpARklingPaper, we created a system to combine the benefits of the haptic feedback of pen- and paper-based handwriting training with the educational benefits of digital learning applications to enhance state-of-the-art handwriting training for children.

3.1 Process

To create SpARklingPaper, we decided to use the iterative user-centered design (UCD) process based on the four phases analysis, design, implementation, and evaluation. We acquired two external advisers/experts to support us and provide domain knowledge. One was a teacher of a nearby elementary school lecturing first- and second-grade pupils. She was available for regular expert interviews and gave us feedback regarding our SpARklingPaper design decisions, and provided insight into the state-of-the-art school curriculum. She further organized a focus group interview with three additional elementary school teachers and provided contact to parents of first- and second-grade pupils for the online survey. The other external advisor was a mother of two children going to elementary school and kindergarten. She became part of our development/authors team and was, therefore, directly involved in the design process and prototype implementation. She ensured that all design decisions were focused on young children and frequently tested our prototypes with her children. This way, we could include two children as users during our UCD process and frequently receive feedback of children. We conducted three iterations of the UCD process to finalize our SpARklingPaper concept as well as a prototype. The first iteration of the UCD process ended after a concept pitch to our expert, the second after the focus group interview, and the third after the user survey. The resulting prototype was the base for our user study.

3.2 Design Goals

We had several requirements for our digital handwriting support system, which were based on the feedback of our UCD process advisors/experts, the elementary school teachers of the focus group, and on previous works (see Section 2). Due to the genuine haptic feedback of pen and paper and its importance for learning handwriting [2, 2, 12, 27, 30, 71] our system should support writing with any pen on arbitrary paper. This way, the children should not have to change their pen when practicing with our system, allowing seamless integration in class and practicing at home without any skill transfer issues [3, 42, 66]. This makes the system further flexible to support a pencil, filler, ball pen, or anything else required by the teachers and, therefore, support all currently standard writing instruments used in handwriting training nationally as well as internationally. For example, in the case of our partner elementary school, this is a triangle-shaped pencil for first- and second-grade pupils.

Literacy is an important issue worldwide [15, 51]. We want to increase handwriting skills with our system for as many children as possible. To ensure that our system would be accessible to a broad audience, we wanted to reuse already commonly available and cheap devices developed for the mass market and have only small extra costs for families or schools.

The system design should be appropriate for use in schools and at home to support lectures, homework, and homeschooling. This means that it should be usable for parents with their children at home as a single instance and also scale to a class setup with multiple simultaneous instances when used at school. To not need two systems, one for school and one for home, a mobile system would reduce costs. It could also be shared between multiple children or families, further increasing accessibility and reducing costs.

During the iterative UCD process, we further added the requirement to create our own learning content easily based on configuration files and without changing the software. This was based on the focus group expert interviews and would allow easy and flexible usage of the prototype for a user study and the teachers an easy possible usage in class.

To guide our development, we transferred these requirements into the following design goals:

(1) To create a system where children can use the pen of their choice to write with digital support on standard paper.
(2) To create a system based on common and cheap hardware that is affordable for everyone.
(3) To create a mobile system usable with multiple simultaneous instances in class and as a single instance for homework or homeschooling.
(4) To provide the possibility to non-developers to create their own learning content for our prototype.

3.3 System Concept

For the system concept of SpARklingPaper, we considered our four design goals. Design goal (1) specifies using standard pens and papers due to the genuine haptic feedback. To build with them a digital education system for children, we propose AR. As typical AR devices such as HMDs (not approved for children [50, 55]) and projectors (occlusion issue [34, 35, 56, 61]) are not suitable for our goals, we decided to put a standard paper on a bright screen that could shine from below through the paper and augment it that way. This would solve the occlusion issue of projectors, as the augmentation is done from below and works without glasses/HMDs why it is suitable for children.

Design goal (2) should ensure that our system would be accessible to a broad audience and usable by as many children as possible. Tablets provide a bright screen and fulfill design goal (2) as it means reuse of mostly already available devices which are further developed for the mass consumer market and, therefore, in an affordable price range for most people. For this, we decided to base our concept on them and put a standard paper on a tablet's touchscreen.

As an integral part of digital educational apps is to provide learning progress feedback, SpARklingPaper needs to recognize what the children have written. Typical handwriting education apps use the tablet's touchscreen to recognize the input of the stylus or the finger, but this would not be possible with our pen- and paper-based approach. As commercial companies also have identified the importance of haptic feedback for digital handwriting, specialized products for tablets exist. For the iPad [37], a so-called "paper-like" protective film is available [28] which tries to simulate the surface of a paper on the touchscreen, especially when used together with an Apple Pencil [36]. Further active styluses exist [44, 70] which all promise a paper-like experience on the tablet's screen. Using such a protective film and a specific stylus would be an easy solution for the input, as the touchscreen could still be used. It would be further compatible with commercial handwriting training apps. However, we have refrained from using them and stuck with our design goal (1), as children would practice as in previous

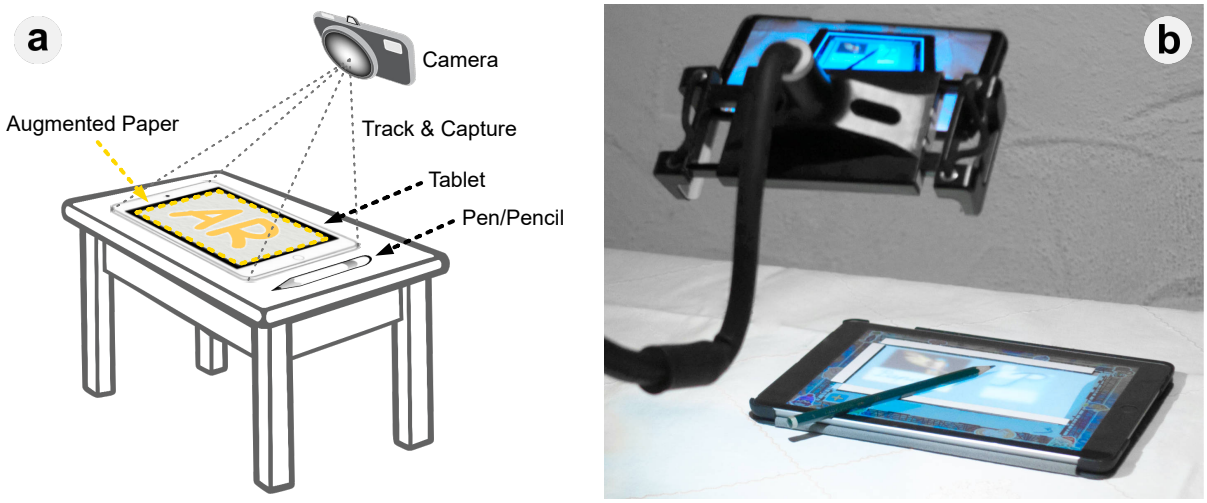

Fig. 2. We present the concept of SpARklingPaper (a), which is based on a pen, a paper, and a tablet. The paper is placed on the tablet's screen. The screen shines through the paper and is augmenting it from below. To capture the pen writing on the paper, a camera is used to track and capture the paper and the written letters. The camera is placed diagonally behind the tablet to not interfere with children during writing. We further present a prototype implementation of our concept (b). It is based on an iPad and an Android smartphone. The paper is fixed on the tablet's screen with magnetic stripes. The smartphone is held by a standard holder. The smartphone is not entirely fit into the holder, and the tablet is facing toward the smartphone for illustration purposes.

works (see [8, 53, 58]) a proprietary device combination, and not the everyday standard pen and paper, which would cause the same skill transfer issues [3, 42, 66] we planned to solve with SpARklingPaper.

As a solution to recognize the pen- and paper-based input, we propose to use a standard camera (see design goal (2)) that tracks and captures from above the paper and the written letters on it. The camera and tablet should be connected so that SpARklingPaper can evaluate the writing success and provide feedback with augmentation on the paper to the children. As the camera should not interfere with the children during writing, we propose positioning it diagonally behind the tablet and directed at the paper. This position has no occlusion issue by the hand of the children, e.g., as AR systems based on projectors. The camera does not always have to observe the writing, but only when a task is finished, and the children have stopped writing, which was important for our experts due to privacy reasons. The output that shows the task and guides the children is provided by the tablet from below the paper and cannot be occluded. Figure 2a shows our system concept.

Using mobile and commonly available devices, we already considered important properties for design goal (3) to create a small and portable system, e.g., unlike systems based on projectors, which is usable in class and at home. We decided to implement a prototype to present a concrete solution considering all design goals, including design goal (4) to easily create own learning content by non-developers.

3.4 SpARklingPaper Prototype

Our SpARklingPaper prototype is based on a pen, a paper, a tablet, and a smartphone as a camera (see Figure 2b), following our previously introduced concept (see Figure 2a). It is an artifact contribution, according to Wobbrock and Kientz [75]. We will describe the details of the iterative implementation of the different components and the final prototype in the following.

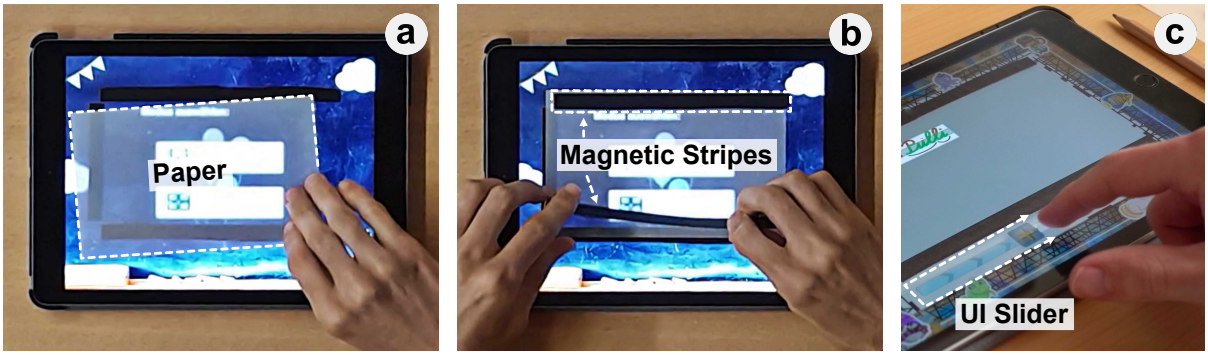

Fig. 3. To augment paper with our prototype, the paper must first be placed accurately on the magnetic stripes (a) before being clipped between them and fixed to the screen (b). Then, when a task is finished and the children want to receive feedback, they can use the slider at the bottom of the screen (c) and slide from left to right to continue.

3.4.1 Tablet-based Paper Augmentation. As proposed in the concept of SpARklingPaper (see Figure 2a), we placed the paper on the touchscreen of a standard tablet. The screen is shining from below through the paper, which allows us to augment it, and it displays the letters, explanations as well as feedback for handwriting training (see Figure 1). We preferred slightly transparent paper with the same friction as standard paper to increase the augmentation effect. Only an exclusively white sheet of paper is required for our prototype since the tablet displays all necessary information, such as letters and lines. This decision was based on the advice of our expert, as it allows very flexible learning units with, e.g., different-sized letters, which is important for design goal (4) to create own learning units. We will describe an example learning unit in detail in Section 3.4.3.

The paper is fixated with magnetic stripes on the tablet so that it does not slide during writing, and the children do not have to hold it permanently. We use magnetic stripes that stick on one side to attach them to the screen and clip the paper by three sides for a tight fit. Figure 3a shows the three magnetic stripes and a paper that is currently placed on them and clipped as well as fixated in Figure 3b with them. Magnetic stripes were found to be the most efficient and convenient way to fix the paper as, e.g., photo corners are more tricky and slow to put the paper in, especially for children, as evaluated with the two children of our expert, not that durable, and hold the paper not that strong. Depending on the child's handedness, we use no stripe on the dominant hand's side so that the hand can be placed comfortably on the screen. The lower stripe is not unpleasant for writing as tested by our experts' children. Figure 1a/b and 3a/b show the stripes placed for a right-handed child. We further use a touchscreen protection film as a precaution so that the stripes are not directly glued on the screen, which also protects it from the pen's pressure. Considering design goal (2), these are all cheap components.

We use a marker to track the position of the tablet and designed it child-friendly with colorful pictures. To display the marker as large as possible, we decided to put it at the screen's outer edges, framing a large enough inner area that augments the paper (see Figure 2a / 4a / 5b). As the paper should not cover the marker, which we tested would significantly reduce recognition, we cut paper pieces fitting to the augmentation area and the magnetic stripes. Tracking the tablet is sufficient, and no calibration of the exact position of the paper or an extra marker on the paper is necessary, as the feedback is calculated based on the stroke position on the paper in relation to the augmentation and the marker.

When finishing the current writing task, the children can trigger the automatic feedback calculation with a slider at the bottom of the screen (see Figure 3c). The tablet will then send a command to the connected smartphone (see Figure 2b) to capture the written letters. We implemented a slider and no button, which prevents

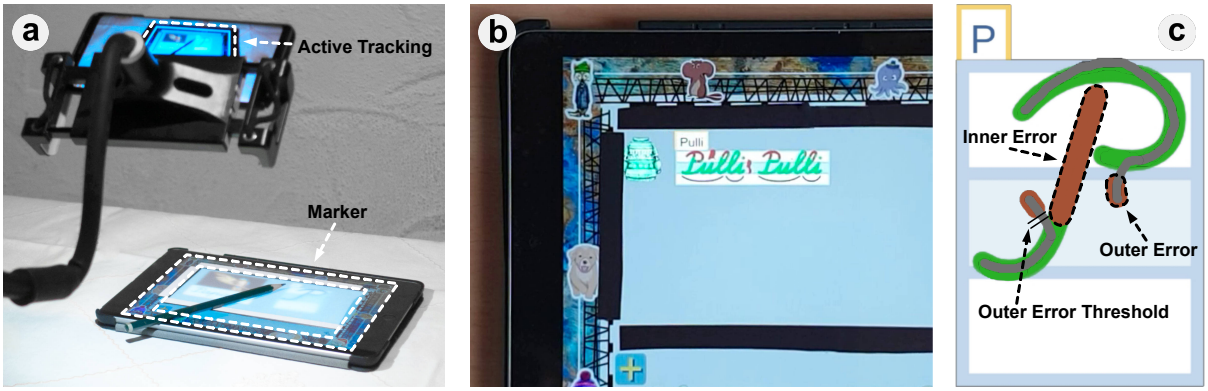

Fig. 4. The smartphone tracks the tablet with a marker (a). When the marker is in the field-of-view and can be tracked from the current position, the smartphone indicates active tracking to the user by showing a white border around the marker. The feedback on the quality of the written letters is provided to the children in green and red. Picture (b) shows this twice for the word "Pulli", which means pullover in German. We distinguish between an inner error and an outer error (c). There is a small threshold for the outer error to ignore small outliers.

unintentional commands while simultaneously being easily usable. We animated the slider further to motivate children to make progress.

3.4.2 Smartphone-based Writing Recognition. We use a smartphone as a camera to recognize the written letters, as the touchscreen cannot detect the pen's writing on the paper. Using a smartphone perfectly fits, as it is a standard, usually already available device (see design goal (2)). Further, modern smartphones have very high-quality cameras and also support WiFi as well as Unity [67] apps out of the box. This allows an easy and powerful integration with the tablet. The smartphone camera has to be placed approximately half a meter above the tablet's screen so that the whole tablet can be captured and the marker, as well as the written letters, can be recognized. To fix the smartphone in this position and consider design goal (2) to use cheap components designed for the mass consumer market, we use an affordable standard smartphone stand compatible with many currently available devices (see Figure 2b). When positioning the smartphone, it is further essential to capture the tablet's screen and choose a position that does not disturb the children while writing. Therefore, we chose a position in our prototype that is diagonally behind the tablet. The smartphone shows the user if the tablet is tracked from the current position (see Figure 4a). To be compliant with data privacy laws as well as the ethics board of the university regarding systems designed for children, our application only takes pictures for a few seconds during calculating feedback and deletes them afterward. Other personal data is not stored. We further advise positioning the smartphone so that only the tablet is captured and, e.g., no faces of the children.

To track the tablet by a marker, we chose Vuforia [69] due to the multi-platform and Unity [67] support. When the tablet marker is in the field-of-view of the camera, it is automatically tracked, and the strokes on the paper can be captured. Nevertheless, during writing, the tablet is occluded by the pen and the child's hand, making tracking and providing feedback difficult. Compliant with our concept, we decided to provide feedback after finishing a coherent task, e.g., writing one line. The children have to put down their pen and use the previously explained slider, and the smartphone can then capture the written strokes without occlusion.

We use OpenCV to cull, transform, and binarize the needed information from the captured picture to determine the drawn pen strokes. We iteratively applied the following steps during development to improve the image quality and optimize stroke recognition. We hide everything from the user interface (UI) besides the needed

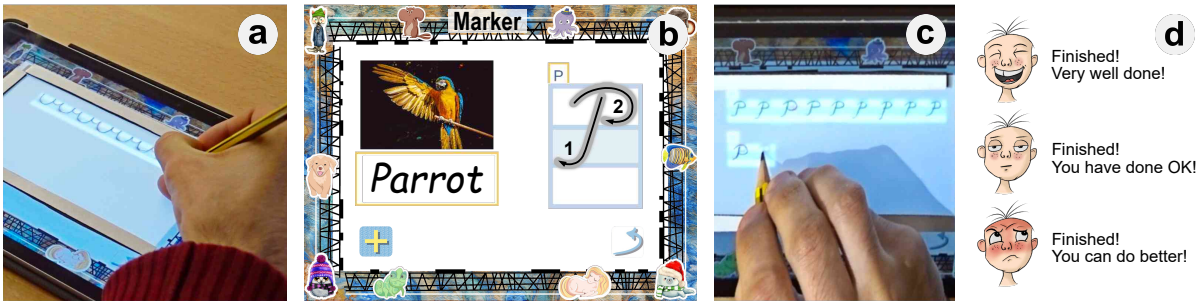

Fig. 5. A typical learning unit starts with swing exercises (a) to activate the graphomotor handwriting skills before a new letter, 'P' in this case, is introduced (b). New letters are introduced first, with an animation showing the stroke sequence and a short sound pronouncing it. Children further see an animal starting with that letter, which is a parrot for 'P'. A further part of a typical learning unit is to practice the letter several times (c). In the end, the children will receive feedback to self-assess how well they have performed (d).

marker during image capturing. Therefore, only the strokes and the paper remain and are captured by the camera, and nothing of the augmentation. A threshold is used to categorize pixels into strokes or not, which is automatically adapted depending on the illumination of the image part inside the marker representing the augmentation area. The image can be separated clearly in written letters and background by concerning its histogram. There is always a significant peak where the white pixels of the background paper are detected. When this peak is cut off, the dark pixels remain, which belong to the pen strokes. We further activate the smartphone's flashlight to improve illumination both during handwriting training as well as image capturing. Therefore, the paper is illuminated from above and below from the tablet screen. To reduce noise, we capture multiple images. By categorizing pixels as strokes, we can detect even thin strokes if at least one taken image returns true as the threshold is exceeded. The binarized images are sent to the tablet, where the feedback calculation is performed. We tested this recognition with multiple light settings, which resulted in solid and sufficient recognition rates.

To calculate if children have written letters correctly, the displayed output and the captured images are compared pixelwise. Correctly written strokes are highlighted in green, whereas incorrectly written strokes are displayed in red (see Figure 4b, Figure 1c/d). We distinguish between an inner error, meaning that a part of a letter is missing, and an outer error, meaning that strokes are outside the letter template (see Figure 4c). This differentiation was added based on the teachers' feedback during the focus group interview, which is a common metric to rate children's writing at school. We implemented a small threshold for the outer error to only detect bigger outliers clearly outside the given letter's bounds and forgive the children minor inaccuracies. As displayed letters are bigger than standard pen strokes, we implemented an algorithm that markers an area of a letter as correct if a stroke is detected inside the bounds of the letter template. This ensures that a letter appears completely in green when all strokes are inside the template, and no inner error exists.

3.4.3 Learning Unit Design. We present an implemented example for the cursive upper case letter 'P' to show what a learning unit with our augmented handwriting system can look like. It is based on the curriculum of first- and second-grade writing education, e.g., see "Jo-Jo Fibel" [14], explained to us by our expert the elementary school teacher and discussed and sophisticated together with the other teachers during the focus group interview (see Section 4.1).

Two screens of swing exercises are shown to get started with the graphomotor writing skill training (see Figure 5a). The upper line is presented first, and when finished, the child uses the slider to get feedback as well as the following line. The following exercise screen is shown when the current screen is complete, and the paper has

to be changed. In the next step, the letter is introduced, with an animation of the writing sequence of the strokes and a short sound pronouncing the letter (see Figure 5b). On the same screen, a picture of an animal whose name starts with the corresponding letter is shown, and a voice-recording with the animal's name is played. This is a parrot in our case. Then, the child writes the letter once in a large size at the position where 'P' was previously introduced with the stroke sequence (see Figure 1b). On the following screen, the letter has to be written multiple times in standard size and is animated for the first three times (see Figure 5c). On the last screen, the children have to write short words with the new letter and letters that they already know (see Figure 4b). Unknown letters can also appear in a word. However, they do not need to be written and are marked in blue. To illustrate the words, images representing them are shown next to text lines (see Figure 4b). At the end of the exercise, the end screen is shown, and the overall score is displayed to the children as a smiley and text feedback (see Figure 5d).

3.4.4 Software Implementation. For the software development, we did consider design goal (3), which states that the system should be usable and scale for a classroom setup or a large roll-out for homeschooling, and used Unity [67], which supports both Android and iOS devices. This way, we were able to create a platform-independent software running in different modes on tablets as well as on smartphones from all major vendors. Our example setup is, e.g., based on an iPad 6 running iOS and a Samsung Galaxy S9+ running Android.

The smartphone and the tablet are connected with a TCP connection via WiFi, where the tablet is acting as a server, and the smartphone is a client. The tablet will send cyclic broadcasting messages during startup containing a configured unique name displayed on the smartphone. If multiple tablets are on the same network, the smartphone displays all, and the corresponding one can be chosen to connect exclusively. Therefore, multiple systems are possible via the same WiFi connection, allowing easy classroom setups (see design goal (3)).

The possible exercises of our SpARklingPaper prototype are pretty flexible and configurable by a JSON file which can be individually authored. It is possible to configure which text or letters should be practiced, including the visualization of syllables as curved lines below the word (see Figure 4b), the font size and block or cursive letter configuration, swing exercises, the letter stroke order animations defined by an SVG file, the images shown on the left side of a letter/line, and sounds played during letter introduction. The human-readable and easy modifiable but still very flexible JSON format is a perfect fit to fulfill our design goal (4).

4 USER SURVEY

As the first step to evaluate our SpARklingPaper concept, we conducted a focus group interview followed by two user surveys.

4.1 Focus Group Interview

Our expert, the teacher at the elementary school, made it possible to show our system in a demo session to three further teachers of the first and second grades and met them for a focus group interview at their school. They gave us two essential hints, which we incorporated into the final system. The first was the additional visualization of the outer error explained earlier (see Figure 4c). The second was to implement an overall score to provide the children feedback about their progress at the end of a learning unit (see Figure 5d).

We also discussed a possible use of SpARklingPaper in class. Therefore, we decided to include an easy solution to create learning content without changing the prototype's software (see design goal (4), the JSON file is explained in Section 3.4.4). This allows to quickly adapt the prototype to the letters and teaching concept that should be learned in class. During the user survey, the teachers were further our participants and established contact with the parents who also participated.

4.2 Survey Design

We created two online surveys, one targeting the teachers of our partner school and one targeting the parents of the children currently learning writing in second-grade. An online survey was compliant with the COVID-19 regulations of the government as well as the university and allowed us to get aware of the children's and parents' needs regarding writing training as well as a first feedback for our SpARklingPaper concept and the prototype implementation, before COVID-19 would allow conducting a user study in-person. Both surveys had similar parts for demographics, children's writing skills before school, commercial writing apps, what motivates the children to learn, and questions regarding our augmented handwriting system SpARklingPaper and its handling. Specific for the parents were questions regarding how their children do their homework, their and their children's technical skills regarding mobile devices, and the current writing lessons at school. Specific for the teachers were questions regarding their technical skills and how to integrate SpARklingPaper into their lessons.

Two videos supported the assessment of SpARklingPaper. The first one introduced our concept and the developed SpARklingPaper prototype. After watching this video, the participants directly answered the corresponding questions. A second video showed the handling and assembly of SpARklingPaper, once again followed by corresponding questions. Our supplementary video is based on the survey videos. The complete survey consisted of 63 questions for the parents and 44 questions for the teachers, which were asked in their native language. Questions were either formulated based on a Likert scale or requested a free-text answer. We provide all questions and the answers, including the summarized free-text answers, as supplementary material.

4.3 Participants

We received 25 responses of parents (age: $29 - 56, M = 40.98, SD = 6.10$) and 4 responses of teachers (age: $27 - 36, M = 32.25, SD = 3.56$). The children of the interviewed parents were between 7-8 years old ($M = 7.32, SD = .47$, 11 female, 14 male). Teachers were compensated with 5€ for participation, and 10 vouchers, each worth 20€, were raffled between the parents.

5 SURVEY RESULTS

The following section shows the results of our surveys. For parents, results are stated with relative frequency (N=25), whereas due to the lower number of participating teachers (N=4), their results are stated as absolute frequencies in brackets. The same applies to the qualitative analysis of the free-text questions, where also absolute frequencies are used to state how many participants have given that answer. When not stated otherwise, Likert scales were used, ranging from 1 (= "very bad") to 5 (= "very good").

5.1 Insights Into Current Literacy/Handwriting Training

The survey started with questions regarding the children's writing skills before school, followed by ones regarding the current school lessons, homework, motivation, and commercial literacy training apps.

5.1.1 Children's Writing Skills Before School. 68% of the parents stated that their child practiced writing before entering elementary school (monthly: 12%; weekly: 48%; daily: 8%), resulting in knowledge of how to write 4-20 letters ($M = 11.64, SD = 4.95$). They were between 4-6 years old ($M = 4.64, SD = .56$) when they were able to write their name. The teachers rated the writing skills of the children at the beginning of first-grade moderate (n=2) and low (n=2).

5.1.2 Insights regarding Current School Lessons, Homework, and Motivation. According to the teachers, the children need 3-7 days to learn a block letter (n=4) and between 1-2 (n=1) and 3-7 days (n=3) to learn a cursive letter. The current literacy training lessons were rated positively by the parents ($M = 4.12, SD = .77$). When asked what could be improved, parents had not much to criticize. They answered that the lessons could have more

variations (n=3), and teachers could provide more individual feedback (n=2). A stated positive fact are the plenty exercises, which should be kept on this level (n=3). 64% of the children need between 10-30 minutes (<10 min: 8%; 30-60 min: 24%; 60-90 min: 4%) for their daily handwriting training homework and practice additionally less than 30 minutes (<10 min: 60%; 10-30 min: 40%). 76% are supported by their parents for more than 10 minutes during homework (<10 min: 24%; 10-30 min: 48%; 30-60 min: 20%; 60-90 min: 8%). While the children are not unmotivated to do homework from the beginning ($M = 1.92, SD = 1.16$; 1 = "strongly disagree" - 5 = "strongly agree": scale used for the following as well), their motivation drops during homework, and they stay unmotivated ($M = 2.56, SD = 1.24$). This results in a nearly even split in children conducting their homework without a break ($M = 2.80, SD = 1.41$) and children needing a break ($M = 2.72, SD = 1.54$). The teachers reported similar but slightly more positive observations for the children's motivation during lessons (see questions 11-14 supplementary material). To keep motivation high, parents usually promise a reward (n=6), e.g., playing with friends after finishing homework, commend their children (n=3), or talk with their children about their goals (n=2) and wishes (n=3), which is quite similar to the teachers approach during lessons (reward (e.g., diligence points): n=2; commendation (by the teacher): n=2; gamification (playing and learning with other children): n=2). New technologies have the potential to remain interesting for children for a very long time ($M = 3.84, SD = .88$).

5.1.3 Commercial Literacy Training Apps. Literacy training apps were not very common, and only 6 parents have used them at home and none of the teachers during their lessons. This is in line with the fact that digital devices are also used only monthly (n=2) or never (n=2) during lessons, which were primarily tablets or projectors. The parents also stated that literacy training apps were not used very often ($M = 2.67, SD = .47$; scale from 1 = "never" to 5 = "very often"). They rated them further only average (usability: $M = 3.33, SD = .94$; learning effect: $M = 2.83, SD = .69$; functionality: $M = 3.50, SD = .50$; overall score: $M = 3.67, SD = .75$).

5.2 SpARklingPaper Augmented Reality Handwriting System Review

The survey continued with questions regarding the SpARklingPaper augmented handwriting system approach, the maturity of our implemented prototype, possible usage of SpARklingPaper at school and at home, as well as technical preconditions.

5.2.1 SpARklingPaper Augmented Handwriting System Approach. After showing a video presenting our system, we asked the parents and teachers questions regarding the combination of a pencil, a paper, and a tablet, the presented functionality, the usability, the helpfulness for learning writing, the feedback mechanism for written letters, the impact of this system on the children's motivation to learn writing, and their overall impression. We specified the used pen as a pencil in the survey, as the children of the parents are all using pencils at the moment to learn handwriting. As the questions were identical for parents and teachers, we present all answers (N=29) in Figure 6 together. It shows the descriptive data on the left (pencil+paper+tablet: $M = 3.41, SD = .85$; functionality: $M = 3.93, SD = .78$; usability: $M = 3.79, SD = .96$; helpfulness: $M = 3.72, SD = .83$; feedback: $M = 4.20, SD = .76$; motivation: $M = 3.86, SD = .63$; overall score: $M = 3.86, SD = .86$), and the same data as diverging stacked bar chart on the right.

When asking them what they like about our system, parents and teachers stated the way we provide feedback (n=9) and its objectiveness (n=4): *"[Negative] feedback by a tablet would be better accepted than by us parents [or teachers]."* They also think that our system will motivate the children (n=7) and is easy to use (n=3): *"Digital things are always more interesting for children than static paper."* That our system, despite using technology, still relies on pencil and paper was also explicitly stated as a benefit (n=3) and that this is a good idea (n=5): *"[I like] that it is possible to work with common tools." "[I like] that paper and tablet are used together." "The idea behind this system is great."* Parents and teachers stated that they think critically of the further screen-time (n=4) and the

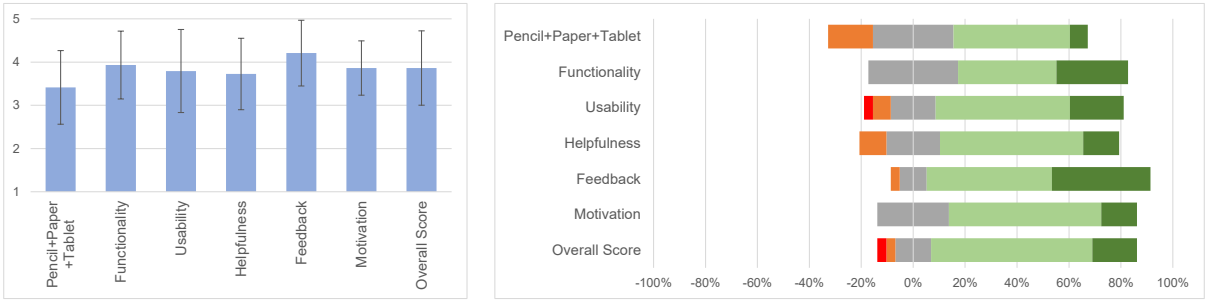

Fig. 6. We asked the parents and teachers (N=29) questions about how they rate our presented system regarding the combination of pencil, paper, and a tablet, the presented functionality, the usability, the helpfulness for learning writing, the feedback mechanism, the impact of this system on the children's motivation to learn writing, and their overall impression. The Likert scale (1 = "very bad" to 5 = "very good") results (mean and standard deviation (SD)) are presented on the left, and the same data as a diverging stacked bar chart on the right.

ergonomic placement of the hand on the tablet (n=4): *"The tablet adds height and provides no plain surface, which seems inconvenient."* One parent suggested using larger tablets for more convenience.

5.2.2 Prototype Maturity. The second video showed details on how to assemble our prototype with the tablet and the smartphone, which is clamped in a holder, and how to use everything. We then asked questions regarding this setup and if parents and teachers think that this could influence handwriting training, if they would assemble and disassemble it every day, if they think they, as well as the children, could handle the assembly, and if children would be able to use the system unsupervised. As the questions were identical for parents and teachers, we present all answers (N=29) in Figure 7 together. It shows the descriptive data on the left (setup influence: $M = 2.41, SD = .85$; daily setup: $M = 2.1, SD = 1.30$; adult assembly: $M = 4.38, SD = .93$; child assembly: $M = 2.69, SD = 1.12$; unsupervised usage: $M = 3.31, SD = 1.09$), and the same data as diverging stacked bar chart on the right.

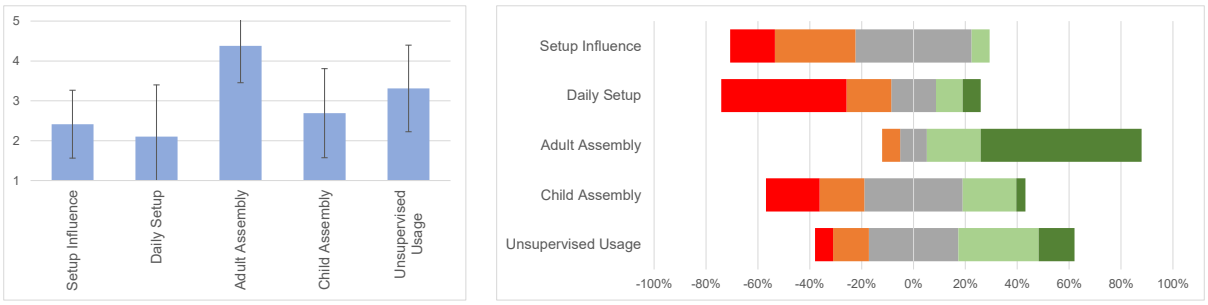

Fig. 7. We asked questions about the setup of our system and if parents and teachers (N=29) think that this could influence handwriting training, if they would assemble and disassemble it every day, if they think they, as well as the children, could handle the assembly, and if children would be able to use the system unsupervised. The Likert scale (setup influence: 1 = "very disturbing" to 5 = "not at all disturbing "; others: 1 = "certainly not" to 5 = "for certain") results (mean and SD) are presented on the left and the same data as a diverging stacked bar chart on the right.

The majority of the parents (92%) and all teachers stated that they would like to spend a maximum of 5 minutes assembling the system. A class-setup should be usable after at least 15 minutes, but preferably much faster.

Parents and teachers (N=29) stated, based on the video, possible concerns about how to assemble the smartphone holder (41%), the WiFi connection between the smartphone and the tablet (55%), the magnetic stripes (17%), and the touchscreen protective film (31%), whereas 14% see no issues. They further would like a detailed manual for the assembly.

5.2.3 SpARklingPaper at School and at Home. We asked the parents if our system could extend the current school lessons, to which 44% agreed (44% undecided; 12% declined) and if our system could improve learning writing at home, especially during COVID-19 pandemic, where 72% agreed (16% undecided; 12% declined). The teachers also stated that the system could be helpful in school (n=3) and at home (n=4) as an extension of the standard exercise sheets (school: n=3; home: n=4) and to introduce new letters (school: n=2; home: n=2). All teachers stated that the app should be similar in its exercises to the standard exercise sheets. Nevertheless, teachers are cautious about using the system during their lessons (no: n=3; maybe: n=1) based on the assembly and handling concerns, but also due to organizational obstacles (e.g., no extra room or WiFi issues at school).

5.2.4 Technical Preconditions. As our system needs some technical preconditions and technical skills, we asked the parents and teacher regarding their mobile device skills ($M = 4.14, SD = .78$) and of their children ($M = 3.24, SD = .99$) as well as parents daily device usage (<1 h: 20%; 1-2 h: 28%; 2-4 h: 44%; 4-6 h: 4%; >6 h:4%) and of their children (<10 min: 24%; 10-30 min: 44%; 30-60 min: 24%; 60-90 min: 4%; >90 min: 4%). In general, parents stated that children handle new technologies and applications well ($M = 3.96, SD = 1.04$). Most of them (76%) own a tablet primarily based on Android (89%) as well as a smartphone (100%), also primarily based on Android (84%). As our system needs some space, we asked if parents have the possibility to install it permanently, which 64% confirmed. Furthermore, they stated that 92% of the children have their own desks. 72% would be willing to prepare their own tablet with, e.g., magnetic stripes or protective film.

5.3 Discussion

Our results show that despite 68% of the children already practicing writing before entering elementary school, their writing skills were still rated relatively low by their teachers at the beginning, which is in line with Danna and Velay [15] and Marquardt et al. [51]. Learning how to write is a long and challenging undertaking, and children need approximately one week per letter. Therefore, learning how to write takes a large part of the first- and second-grades showing the huge potential of literacy/handwriting training systems or apps. Nevertheless, commercial apps are still not used very often, despite their previously explained benefits (see Section 2), which raises the question of how a system has to be designed to be frequently used. This highlights the importance of our work.

Teachers and parents confirmed the importance of pen and paper's genuine analog haptic feedback for handwriting training and the advantages of the augmented feedback of SpARklingPaper for class and homework. The implemented functions in our prototype were rated positively as well as the usability which shows that our prototype is ready for an in-person user study. We will discuss the further findings of the survey together with the results of the user study in Section 8.

6 USER STUDY

We conducted a user study to answer our RQ and created two hypotheses for evaluation.

6.1 Hypotheses

The user survey showed that our SpARklingPaper prototype already has the fidelity for a user study and to answer our RQ *"Does the combination of a pen, a paper, and a tablet for writing training, as SpARklingPaper provides it, impact pen-handling, assumed writing training-success, usability, motivation, feedback understandably, perceived*

functionality, and the overall system impression, compared to a state-of-the-art stylus- and tablet-based training?". Based on our stated related work (see Section 2), our assumptions to design SpARklingPaper (see Section 1 and 3.3), and the survey results, we formulated two hypotheses:

H1: *"SpARklingPaper improves (1) pen-handling, (2) assumed writing training-success, and provides (3) an overall higher impression compared to state-of-the-art stylus- and tablet-based learning."*

H2: *"SpARklingPaper does not decrease (1) the usability, (2) the motivation, (3) the feedback's understandably, and (4) the perceived functionality compared to state-of-the-art stylus- and tablet-based learning."*

H1 states our intention that SpARklingPaper can combine the genuine haptic feedback of analog pen and paper with the digital support of apps, resulting in a superior learning experience. H2 takes into account that SpARklingPaper needs more hardware than state-of-the-art stylus- and tablet-based apps. As the survey implies that this could have negative effects, we want to investigate this further but hypothesize that this overhead will be accepted.

6.2 Study Design

To evaluate our SpARklingPaper design, our RQ, and the hypotheses, we conducted a within-subject user study and compared it against a state-of-the-art stylus-based condition. For the stylus condition, we used the same app as SpARklingPaper on a similar iPad but activated the touchscreen-mode to be stylus compatible. We used an Apple Pencil [36] as well as the Paperlike film [28] to improve the friction of the touchscreen and to use the currently most advanced commercially available hardware for our baseline condition. The learning unit (see Section 3.4.3) was the same for both conditions and taught the letter 'P'.

To assess our hypotheses, we created Likert scale-based questionnaires for both the parents and their children covering the pen-handling, the assumed writing training-success, the usability, the children's motivation, the provided feedback's understandability, the perceived functionality of our app, and the overall system impression. We further created questions for semi-structured interviews covering these topics to get a deeper understanding by a qualitative analysis. Besides these questionnaires and interview questions following each condition, we further created a follow-up questionnaire directly comparing both conditions and a concluding interview. All questionnaires and interview questions are provided in the supplementary material.

6.3 Procedure

We recruited our participants through convenience sampling and assigned the first condition (SpARklingPaper or Stylus) based on a Latin square. Participants initially received a short non-priming introduction about the topic of the study as well as the prototypes for the conditions, and demographics were collected before the first condition started, where children conducted swing exercises and wrote the letter 'P' (learning unit explained in detail in Section 3.4.3). After finishing the learning unit, the children and the parents answered the questionnaires followed by the semi-structured interview with the parents. Then the second condition started, followed once again by questionnaires and an interview. After both conditions were finished, children and parents answered comparing overall questionnaires followed by an overall interview. The conduction of the learning unit with both conditions was video- and the interviews were audio-recorded for later analysis. Furthermore, notes were taken. The children were compensated with a certificate of participation and a little present. The parents received 20€. Due to the COVID-19 pandemic, the study was conducted based on a hygiene concept approved by the university, which caused no limitations regarding our study design.

6.4 Participants

We collected data from 13 children and 13 parents (N=26). The children were between 4-8 years old ($M = 6.54$, $SD = .97$, 4 female, 9 male, 2 left-handed). The participating parents were between 34-44 years old ($M = 38.15$, $SD = 3.44$,

12 female, 1 male). The children knew how to write between 5-26/all letters ($M = 21.69, SD = 5.81$) and attended pre-school (n=1), first grade (n=10), or second grade (n=2) of elementary school. Most children used a triangle-shaped pencil designed for writing beginners (n=10), while the others used an ordinary pencil.

7 RESULTS

In the following, we will show the results of our user study. For the quantitative results, when not stated otherwise, Likert scales were used, ranging for the children from 1 (= "bad") to 3 (= "good") and for the parents from 1 (= "very bad") to 5 (= "very good"). The interviews were qualitatively evaluated based on a thematic analysis [9, 10].

7.1 Quantitative Results

We will first present the questionnaire results of the children for both conditions, which are then followed by the results of the parents. Due to the inaccuracies of significance tests for samples of similar sizes to ours (N=13 for parents as well as children), we refrained from conducting them (see Field [25]) and report descriptive statistics using standard bar charts as well as diverging stacked bar charts, as higher means combined with lower SDs also highlight relevant differences. The average duration of a session was $M = 87.69$ minutes ($SD = 7.80$) with $M = 22.50$ minutes ($SD = 5.41$) for SpARklingPaper and $M = 17.77$ minutes ($SD = 3.57$) for the stylus condition.

7.1.1 Children's Opinion. After finishing the learning unit of a condition, we asked the children using Likert scales (1 = "bad" to 3 = "good") how they liked each condition (liked system:= AR: $MD = 3.00, SD = .00$; Stylus: $MD = 2.92, SD = .28$) and how they liked to use their own pencil in combination with the paper and the tablet (liked pen:= AR: $MD = 3.00, SD = .00$) as well as the stylus on the paper-like film (liked pen:= Stylus: $MD = 2.77, SD = .44$). We further asked them how they would like to use each condition more often (use more often:= AR: $MD = 2.85, SD = .38$; Stylus: $MD = 2.46, SD = .78$) as well as in school (use at school:= AR: $MD = 2.85, SD = .38$; Stylus: $MD = 2.46, SD = .88$). They also rated how it was to use each condition compared to standard state-of-the-art (SOTA) writing training with pen and paper (compared to SOTA:= AR: $MD = 2.77, SD = .44$; Stylus: $MD = 2.00, SD = .91$). For the SpARklingPaper condition, we further asked usability questions regarding the handling of changing a paper (AR paper change: $MD = 2.69, SD = .48$) as well as how visible the learning unit (e.g., the letters) were through the paper (AR letter visibility: $MD = 2.46, SD = .52$). These values are shown in Figure 8. At the end of the study, we asked the children which condition they preferred. Eight answered SpARklingPaper, three the stylus with the paper-like film, and two were undecided (see Fig. 9).

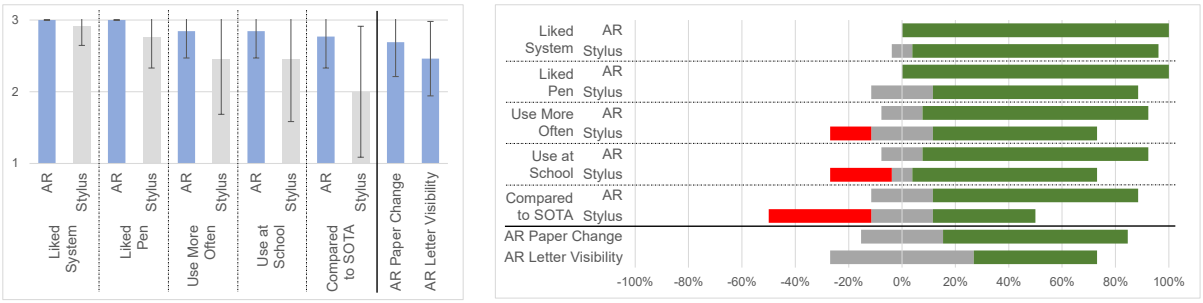

Fig. 8. We asked the children questions about how they liked the conditions/systems, how they liked the pen (own pencil vs. stylus), if they would like to use the conditions more often as well as at school, how the conditions were compared to SOTA paper- and pencil-based writing and questions regarding the usability of the SpARklingPaper system. The Likert scale (1 = "bad" to 3 = "good") results (mean and SD) are presented on the left and the same data as a diverging stacked bar chart on the right.

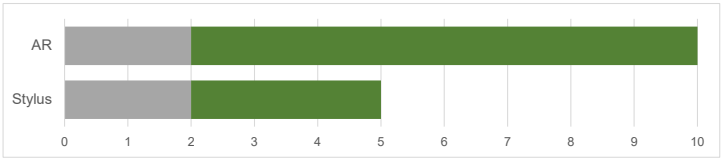

Fig. 9. At the end of the study, we asked the children which condition they preferred. Eight answered SpARklingPaper, three the stylus with the paper-like film, and two were undecided (visualized in grey).

7.1.2 Parents' Opinion. After finishing the learning unit of each condition, we asked the parents using Likert scales (1 = "very bad" to 5 = "very good") how their children mastered the condition. The questions were similar to those of their children as well as the survey (see Sections 5.2.1, 6.2, and 7.1.1) and about how their children liked each condition (liked system:= AR: $MD = 4.23, SD = .83$; Stylus: $MD = 4.00, SD = .82$) and how their children were able to write using their own pencil in combination with the paper and the tablet (liked pen:= AR: $MD = 4.31, SD = .63$) as well as the stylus on the paper-like film (liked pen:= Stylus: $MD = 3.31, SD = 1.11$). We further asked them how they would like it if their children could use each condition more often (use more often:= AR: $MD = 4.23, SD = .93$; Stylus: $MD = 3.69, SD = 1.11$) as well as in school (use at school:= AR: $MD = 4.23, SD = .83$; Stylus: $MD = 3.69, SD = 1.11$). They also rated how they perceived each condition compared to standard writing training with pen and paper of their children (compared to SOTA:= AR: $MD = 3.77, SD = .73$; Stylus: $MD = 3.15, SD = .90$). For the SpARklingPaper condition, we further asked usability questions regarding their children's handling of changing a paper (AR paper change: $MD = 3.69, SD = .95$) as well as how visible the learning unit (e.g., the letters) were through the paper for their children (AR letter visibility: $MD = 3.54, SD = .97$). These values are shown in Figure 10.

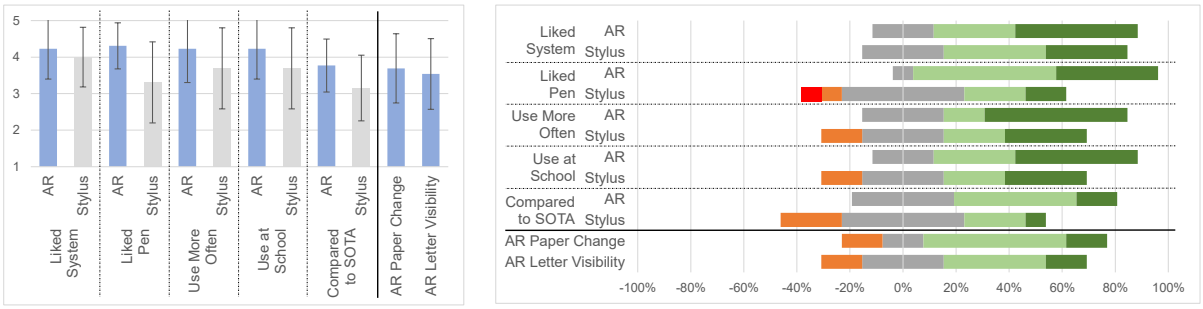

Fig. 10. We asked the parents the same questions as to their children (see Fig. 8). They should consider how their children mastered the conditions and answer them from their parents' perspective. The questions were about how their children liked the conditions/systems, how they liked and could handle the pen (own pencil vs. stylus), if they would like their children to use the condition more often as well as at school, how the conditions were compared to state-of-the-art paper- and pencil-based writing for their children and questions regarding the usability of the SpARklingPaper system. The Likert scale (1 = "very bad" to 5 = "very good") results (mean and SD) are presented on the left, and the same data as a diverging stacked bar chart on the right.

At the end of the study, we asked the parents comparing overall questions similar to the ones in the survey (see Section 5.2.1). We asked them how they rate the shown functionality of each condition for writing training (functionality:= AR: $MD = 4.15, SD = .80$; Stylus: $MD = 4.08, SD = 0.86$), how they rate the handling and the usability (usability:= AR: $MD = 4.00, SD = .71$; Stylus: $MD = 4.54, SD = .66$), the perceived capability of

each condition to support their children in learning writing (helpfulness:= AR: $MD = 4.54, SD = .52$; Stylus: $MD = 3.23, SD = 1.36$), the way of providing feedback (feedback:= AR: $MD = 4.31, SD = .63$; Stylus: $MD = 4.46, SD = .66$), the assumed motivation of their children to use the condition for writing training (motivation:= AR: $MD = 4.62, SD = .65$; Stylus: $MD = 4.08, SD = 1.38$) and if the setup/construction was disturbing during writing (setup influence:= AR: $MD = 4.62, SD = .87$; Stylus: $MD = 4.77, SD = .44$). They should also give an overall rating for each system (overall score:= AR: $MD = 4.15, SD = .69$; Stylus: $MD = 4, SD = 1.15$). These values are shown in Figure 11.

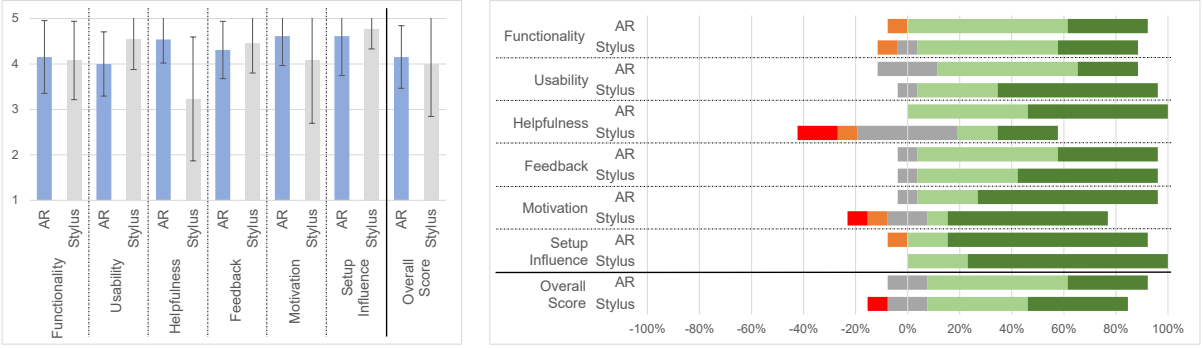

Fig. 11. We asked the parents at the end comparing overall questions for both conditions regarding the functionality, usability, and helpfulness for learning writing, the writing feedback capabilities, the motivational influence on their children, and the influence of the setup on learning. They further gave an overall score for both conditions. The Likert scale (setup influence: 1 = "very disturbing" to 5 = "not at all disturbing"; others: 1 = "very bad" to 5 = "very good") results (mean and SD) are presented on the left and the same data as a diverging stacked bar chart on the right.

7.2 Qualitative Results

After each condition and at the end, we conducted semi-structured interviews with the parents. The interviews were conducted in the participants' native language and were on average $M = 6.67$ minutes ($SD = 4.73$) for the SpARklingPaper condition, $M = 5.81$ minutes ($SD = 2.41$) for the stylus condition, and $M = 1.85$ minutes ($SD = .62$) for the overall feedback. Based on the taken notes, which were completed afterward using the audio and video recordings (see Section 6.2), one author conducted a reflexive inductive thematic analysis similar to Braun and Clarke [9, 10], which needs no coding framework and can be conducted by a single person. Based on their six phases [9] the following themes emerged *(1) Influence of the Pen on Writing Training*, *(2) Device Handling*, and *(3) Media Supported Learning*. We included multiple translated participant statements for good quality [9, 10].

7.2.1 Theme 1: Influence of the Pen on Writing Training. All parents stated that their children could write reasonably well using their own pencil in the SpARklingPaper condition. They were familiar with their own pencil, and the hand posture was natural. *P2* mentioned that it was *"fun"*, and *P4* and *P12* that *"writing was fast [...] and less exhausting [compared to the stylus]"*. *P5* stated that the tablet as a writing pad did not make a difference *"It was like ordinary writing, [the paper was just lying] on a tablet."*. *P6*, *P7*, and *P13* emphasized that SpARklingPaper benefits from training with the target devices pen and paper, which are still primarily used for handwriting in daily business. Despite the overall positive feedback, *P2* and *P10* mentioned that standard paper still has a more familiar feeling than our device setup. As *P1*, *P10*, and *P12* stated, this could be due to the height and edge of the tablet, which should be reduced to improve writing comfort.

The stylus received mixed reviews. *P2*, *P4*, *P6*, *P7*, *P9*, and *P11* mentioned that their children held the stylus unnaturally. *P4* said that *"[He/She held the stylus] clenched and used high pressure ... which is [not the case] with triangle-shaped pencils."*. *P6* stated that *"The [stylus] has low friction [on the tablet screen]."* and *P11* that *"[...] there were huge differences [to standard pencil and paper writing]. He/She used too much pressure and the stylus got slippery [due to sweaty hands] which resulted in an unusual pen handling."*. Nevertheless, some children could handle the stylus well after a short acclimation phase (e.g., *P3*, *P12*). *P8* stated that their child could write similarly well with either their own pencil or the stylus, and *P5* said that the stylus is quite natural but should have a better grip. Most parents (*P2*, *P3*, *P4*, *P6*, *P7*, *P9*, *P11*, *P12*, *P13*) noted that an improved and more ergonomic stylus design that is similar to a triangle-shaped pencil and dedicated to writing beginners could help to improve its handling.

7.2.2 Theme 2: Device Handling. All parents stated that their children could handle both conditions quite well. However, there were minor issues regarding a slipping paper in the SpARklingPaper condition (*P2*, *P3*, *P4*, *P5*, *P8*, *P10*) as well as some input issues with the stylus caused by the hand lying on the active touchscreen (*P1*, *P2*, *P3*, *P4*, *P6*, *P7*, *P8*, *P10*, *P12*, *P13*). Still, none of these were critical to mastering the learning unit successfully.

While, as previously stated, the natural friction of the paper in combination with the pencil was perceived positively, the handling of the paper could be improved due to our participants. *P1*, *P3*, *P5*, *P8*, *P10*, *P11*, *P12*, and *P13* stated that the frequently necessary paper change added a lot of overhead time (*P10*: *"Changing the paper is very time-consuming."*; *P12*: *"[Proceeding] takes more time [than in the stylus condition]."*) and that a larger paper should be used that covers a greater area of the tablet screen to reduce changing frequency (*P6*, *P8*, *P9*, *P12*). The necessary device setup with the additional smartphone on a stand for SpARklingPaper was further seen as critical for a regular usage at home as well as in school, which should be improved for a consumer version (*P2*, *P3*, *P4*, *P6*, *P8*, *P9*, *P10*, *P11*, *P12*). For example, *P3* stated, *"The smartphone holder needs too much space [...] and it can fall or be pushed over by other children [when used in school]."*

7.2.3 Theme 3: Media Supported Learning. Both conditions were well received by all of the children, and the technical devices caused high motivation. As technical devices can cause such an increase in the child's motivation, *P2*, *P3*, *P10*, *P11*, *P12*, and *P13* deliberately stated that their children frequently use the commercially available Anton app [62] to train writing with their fingers (see Section 2.1). *P3*, *P4*, *P5*, *P6*, *P7*, *P8*, *P11*, and *P12* appreciated that technical systems and devices could permanently provide feedback for the children in an objective way, other than parents and teachers can (e.g., *P11*: *"The [automatically] provided feedback is very good and saves the teachers and the parents time [in advising the children]."*) and *P1*, *P5*, *P9*, *P10*, and *P11* would like it to have more detailed feedback as well as gamification to increase the motivation of the children even further (e.g., *P10*: *"The [green/red] feedback should be more detailed [and analyze the shape of a letter] ... and not only if [the stroke] is in the correct area."*). Although these positive aspects of media supported learning, *P1*, *P3*, *P6*, *P7*, and *P13* also emphasized that digital screen-based learning should be controlled and limited in time and should not completely replace traditional teaching and learning. However, *P3* and *P12* explicitly stated that digital teaching and learning could reduce paper consumption which is beneficial for the environment and a disadvantage of our SpARklingPaper system, in their opinion.

8 DISCUSSION

In this section, we will discuss the results of our user study (see Section 7) and combine them with the results of our user survey (see Section 5). We will discuss the benefits of the SpARklingPaper concept, its usability, and the achievement of our design goals.

8.1 Benefits of the SpARklingPaper Concept

We hypothesized that H1: *"SpARklingPaper improves (1) pen-handling, (2) assumed writing training-success, and provides (3) an overall higher impression compared to state-of-the-art stylus- and tablet-based learning."* The survey showed (see Section 5.3) that our SpARklingPaper concept was perceived well based on the video. This can be confirmed by the user study as well. The descriptive data shows the trend that both, children and parents, liked SpARklingPaper very much and even slightly better than the state-of-the-art stylus baseline condition. It was also the overall preferred system considering everything (see Figures 8/9/10/11), and it was rated superior than the classic analog pen- and paper-based writing. These findings show that our concept to combine the genuine haptic feedback of analog pen and paper with the digital support of apps worked out.

The pen-handling of the own pen was also perceived better than the stylus by both children and parents. While the own pen was common and writing less exhausting, participants stated that the stylus shape was not optimized for writing beginners, provided low friction on the tablet even in combination with the Paperlike [28] film, and it got slippery (see Section 7.2.1), resulting in clenched handling. This confirms previous results [2, 12, 27, 30, 41, 53, 71] regarding the low friction of touchscreens for stylus-based writing, especially when compared to our SpARklingPaper approach using pen and paper. Perhaps improved Paperlike films could increase the surface friction and enhance ergonomics for the stylus. Specific wrappers for the Apple Pencil already exist to improve its shape [22], and parents stated their potential. Still, they all try to imitate standard pens, so we argue to use SpARklingPaper, which allows the use of an original pen and not just a stylus mimicking it.

The parents further stated that SpARklingPaper has the perceived higher capability to support their children in learning writing. Therefore, children and parents would like to use it more often at home and in school compared to the baseline stylus condition (see Figures 8/10). This is similar to the results of our survey. SpARklingPaper further practices, different than the proprietary stylus baseline, with the target devices pen and paper. This is why the parents see no skill transfer issue (see [3, 42, 66]). These findings support our hypothesis H1.

8.2 Usability of the SpARklingPaper Concept

We further hypothesized H2: *"SpARklingPaper does not decrease (1) the usability, (2) the motivation, (3) the feedback's understandably, and (4) the perceived functionality compared to state-of-the-art stylus- and tablet-based learning."* Previous works have shown that digital devices can support children's motivation for handwriting training [11, 26, 52, 57, 79]. This is why multiple parents positively stated their children's use of the Anton app [62] for writing training. However, such apps were still used by the minority of both samples (survey and study) and should be improved regarding their learning potential (see Section 5.1.3). The survey already showed this potential of SpARklingPaper not just based on simply using technology, but due to the highly motivating design of our implemented application, which frequently uses feedback and commendation also often used by parents and teachers themselves. The user study confirmed this, and the parents rated the motivation potential as very high and higher than the stylus baseline (see Figures 11). This is very important as the survey showed that children are difficult to motivate and quickly lose the fun in doing homework, especially at home. As children typically practice writing between 10-60 minutes daily, SpARklingPaper can solve a widespread need and reduce the children's needed daily support provided by the parents.

The way how both conditions provided feedback was rated very positively, which shows that the camera-based feedback (recognition of writing) of SpARklingPaper functioned on the same level as the touchscreen-based. Parents of the survey and the user study rated our color-coded feedback as effective because it is individual and provided in a much higher frequency than a teacher can do during class or a parent at home (see Bonneton-Botté et al. [8]), which was one of the few points of criticism of current lessons at school. Apps are further an objective instance for the children, and feedback would be better accepted according to the parents than from themselves. Nevertheless, the feedback could be more detailed regarding letter shape, and gamification could be improved

(see Browne et al. [11] and Martens et al. [52]). The provided feedback was rated slightly higher in the stylus condition. This could be due to the visibility, which was slightly reduced for SpARklingPaper as it was always displayed through a paper which reduced brightness. It could be an improvement for SpARklingPaper to use even brighter displays, as stated by the participants, but both children and parents rated the current brightness as sufficient (see Figures 8/10).

The parents rated the usability of SpARklingPaper nearly equal to the stylus baseline condition, which we interpret as an excellent result due to the more complex setup with an additional smartphone and a smartphone holder and thinking of the precautions participants of the video-based online survey had regarding assembling it. The setups' influence on the children during writing was also rated nearly equal to the baseline, and parents stated that they see no disturbing effect. However, participants stated desired improvements to ease daily usage, such as improved paper fixation, larger paper size, and a more miniature smartphone holder. Based on the interviews, we think that a commercial version of SpARklingPaper should consist of a few small components that can be easily and detachably clipped to the tablet. Nevertheless, our SpARklingPaper prototype is already a much smaller and more mobile system than previous approaches such as those Winkler et al. [74] or Kim et al. [43] provided (see Section 2.2).

Overall, the current prototype of SpARklingPaper was rated as complete and providing sufficient functionality to help their children in learning writing. These findings support our hypothesis H2, and SpARklingPaper could even positively influence the children's motivation compared to the baseline, which shows that parents and children accept a more complex setup for an improved learning experience.

8.3 Achievement of our Design Goals

Besides our hypotheses, we defined four design goals when starting our work (see Section 1). We achieved design goal (1) that children can use the pen of their choice to write with digital support on standard paper when we designed our SpARklingPaper system concept, proved that this concept could be implemented with our prototype, and successfully evaluated it in a user study (N=26). By building our prototype on standard hardware, naming tablets and smartphones, we could fulfill design goal (2) to use common and cheap hardware affordable for most people. These devices are mobile and already frequently used at home and in school (see Section 5.2.4), and our prototype needs no permanent installation, which ensures that we also fulfilled design goal (3) also confirmed by our survey (see Section 5.2.3). By providing the possibility to author learning units based on the human-readable and easy modifiable but still very flexible JSON format, we could also fulfill design goal (4) and enable customization of SpARklingPaper for non-developers such as parents and teachers.

8.4 Summary

We can summarize that we were able to answer our RQ, support our H1 and H2, and achieve our four design goals with our SpARklingPaper prototype based on the results of a user survey (N=29) and an in-person user study (N=26). We want to highlight the importance of the in-person user study with participating children and parents, as SpARklingPaper was rated much better after testing it in real than in the survey (compare Figures 6/7/11). We have shown that using the own pen compared to a stylus is rated higher by parents and children for learning writing after evaluating the system. Future systems should consider this and either improve the stylus on touchscreen writing experience to become more natural or, as we would advise, create systems such as SpARklingPaper, that support actual pen and paper. Concluding, we want to emphasize that SpARklingPaper is a valuable approach to improving the current state-of-the-art of writing training for children that could be used as a blueprint for future systems.

9 LIMITATIONS

Our experts choice and our survey have potential limitations. We recruited the experts, teachers, and parents from one school, limiting them to one nation and one character set, which could have influenced the results. We were also limited to the four first/second-grade teachers at this school who could participate. The online survey was further based on videos, and the participants could not interact with the system to rate it. Also, despite targeting children, it was only possible to receive the feedback of parents and teachers with the survey and how they think that children would interact with the system. Still, we could not receive feedback from the children themselves.

To overcome several of these limitations, we conducted the in-person user study. However, due to COVID-19, this was still limited to convenience sampling and 13 children as well as 13 parents. This is why we refrained from conducting significance tests due to their inaccuracies for our sample size and based our analysis on descriptive statistics (see Field [25]) as explained in Section 7.1.

One factor that could have limited the usability of the stylus is its cylindrical form factor. Compared to the often triangle-shaped pencils of the children, this form provided lower friction and was more difficult to use according to the children and their parents. Specific wrappers exist to change the shape of the stylus to mimic standard pencils [22]. As their usability was not evaluated previously, and we refrained from a third or potential forth condition in our study design, future work should evaluate their usability compared to SpARklingPaper as well as the default stylus form factor.

10 FUTURE WORK

With the feedback of the online survey and the in-person user study, we could think of a commercial design for SpARklingPaper and a user study in class at our partner elementary school.

10.1 Improving SpARklingPaper

The general system approach and the prototypical implementation were rated very positively. Nevertheless, we would improve and support the handling and assembly of the system, as our participants had concerns about this, especially for a commercial version. The paper should be attached differently to the screen than with magnetic stripes. It should be possible to detachable clip this paper-holder to the tablet, and it should support sheets of paper covering a larger area of the screen to reduce the paper change frequency. Furthermore, the smartphone holder should be smaller and ideally placed directly attached diagonally behind the tablet. However, the best case would be that the smartphone is no longer needed as an external camera as the touchscreen itself can directly sense the writing on the paper, e.g., by using integrated optical sensors [54]. The results of SpARklingPaper are a good argument for future tablets to support this out of the box.

Our findings revealed further improvements for handwriting training, such as an observation and correction of the hand posture of the children during writing. Furthermore, we could think of a stroke order detection and correction as well as an evaluation of the used pressure to write with the pen (see Section 7.2). SpARklingPaper, with its optical handwriting detection, has the potential to detect these measures in future versions as well and help handwriting learners to improve their writing even further.

As previously stated, we provide an authoring capability for the learning content based on JSON files. In the future, we could think of lowering the authoring effort by creating a UI-based Wizard to create the content without the need to edit JSON files. This would increase the user group even further.

10.2 Classroom User Study

The current prototype has high enough fidelity to evaluate it further, and we could think of an in-class user study and already discussed a study design with our partnering elementary school. We plan a three weeks user study

at school to compare SpARklingPaper against a stylus-based version of our app, similar to the one used in our current study, and the baseline condition pen and paper currently used in class. We created a within-subject design based on a power analysis and designed together with the teachers appropriate pre- and post-knowledge tests as well as learning units for the three weeks. The conditions should change each week. This way, we will be able to investigate the expected benefits of SpARklingPaper and quantitatively measure the learning outcome itself.

To conduct this study, we have already obtained consent from several authorities. We convinced the teachers from our focus group as well as the principal of the elementary school to allow our study and to conduct it in class for three weeks in the three second-grade classes of the focus group teachers. Having three classes perfectly fits our study design with its three conditions. The teachers will support us in conducting the study during class. We also approved our study with the university officials and with the responsible school authority of the elementary school. We further can borrow 60 tablets from the school and the university and have all the other hardware available to support all three classes with the necessary equipment.

Nevertheless, we are currently on hold with our study plans due to COVID-19. We do not know when this situation will change as there is still no vaccine for elementary school children. Therefore, we plan to wait until the situation improves and authorities and the situation in class allow the study as planned.

11 CONCLUSION

Although digital devices are omnipresent and educational apps provide benefits for learning, they are seldom used for handwriting training, where still pen and paper are the standards due to the genuine haptic feedback. We created SpARklingPaper, an augmented handwriting system that combines a pen, a standard paper, and a tablet and, therefore, their specific benefits the genuine haptic feedback and individual digital learning. The paper lies on top of the tablet and is augmented from below by the tablet's screen to guide the children while learning and to provide the writing feedback captured by a smartphone. Our system design and the prototypical implementation were guided by four design goals developed together with experts, which were teachers and a mother of elementary school and kindergarten children. Their focus lies on comfortable pen writing for children and high accessibility for a vast population by using already available smartphones and tablets, as well as the applicability in class and at home.

Our results show that parents assume that the generally low writing skills of children when starting school, as well as their motivation in writing training, should and can be improved by SpARklingPaper. Children and parents acknowledged in the user study our design concept and rated it higher regarding pen-handling, assumed writing training-success, and overall impression compared to a stylus and tablet baseline condition. The parents positively perceived the objective feedback regarding writing success, and they stated that this has advantages for homework and class. Although SpARklingPaper is more complex to set up, the usability was not negatively affected, and parents and children accepted this for an improved learning experience.

Summing up, SpARklingPaper provides a profound base for high-fidelity haptic feedback handwriting training based on a pen, a paper, and a tablet which could be used as a blueprint for future systems.

ACKNOWLEDGMENTS

This work was conducted within the project ”AuCity 2 - AR in University Teaching Using the Example of Civil Engineering“ and within the project "AuCity 3 - Collaborative and Adaptive MR in Higher Education Teaching Using Civil Engineering as an Example“ both funded by the Federal Ministry of Education and Research (BMBF).

We would like to express our special thanks to all who have supported us. These are the teachers who frequently advised us, the collaborating school as well as the school authority. We would also like to thank all parents and children who participated in the online survey as well as the user study.

REFERENCES

[1] Age of Learning, Inc. 2021. *ABCmouse.com.* https://apps.apple.com/us/app/abcmouse-com/id586328581 (Accessed on 08/12/2021).

[2] Denis Alamargot and Marie-France Morin. 2015. Does handwriting on a tablet screen affect students' graphomotor execution? A comparison between Grades Two and Nine. *Human Movement Science* 44 (2015), 32–41. https://doi.org/10.1016/j.humov.2015.08.011

[3] John R Anderson and Mark K Singley. 1993. The identical elements theory of transfer. *Rules of the mind* (1993), 183–204.

[4] Anoto. 2021. *Anoto Digital Pen.* https://www.anoto.com/cases/anoto-digital-pen/ (Accessed on 08/13/2021).

[5] Rahul Arora, Rubaiat Habib Kazi, Tovi Grossman, George Fitzmaurice, and Karan Singh. 2018. SymbiosisSketch: Combining 2D & 3D Sketching for Designing Detailed 3D Objects in Situ. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3173574.3173759

[6] Rahul Arora, Rubaiat Habib Kazi, Fraser Anderson, Tovi Grossman, Karan Singh, and George Fitzmaurice. 2017. Experimental Evaluation of Sketching on Surfaces in VR. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 5643–5654. https://doi.org/10.1145/3025453.3025474

[7] Michael Bogorad. 2021. *iTrace — handwriting for kids.* https://apps.apple.com/us/app/itrace-handwriting-for-kids/id583678381 (Accessed on 08/12/2021).

[8] Nathalie Bonneton-Botté, Sylvain Fleury, Nathalie Girard, Maëlys [Le Magadou], Anthony Cherbonnier, Mickaël Renault, Eric Anquetil, and Eric Jamet. 2020. Can tablet apps support the learning of handwriting? An investigation of learning outcomes in kindergarten classroom. *Computers & Education* 151 (2020), 103831. https://doi.org/10.1016/j.compedu.2020.103831

[9] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. https://doi.org/10.1191/1478088706qp063oa

[10] Virginia Braun and Victoria Clarke. 2020. One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology* 0, 0 (2020), 1–25. https://doi.org/10.1080/14780887.2020.1769238

[11] Kevin Browne, Christopher Anand, and Elizabeth Gosse. 2014. Gamification and serious game approaches for adult literacy tablet software. *Entertainment Computing* 5, 3 (2014), 135–146. https://doi.org/10.1016/j.entcom.2014.04.003

[12] AHS Chan and PSK Lee. 2005. Effects of different task factors on speed and preferences in Chinese handwriting. *Ergonomics* 48, 1 (2005), 38–54. https://doi.org/10.1080/00140130412331303902 arXiv:https://doi.org/10.1080/00140130412331303902 PMID: 15764305.

[13] Youngjun Cho, Andrea Bianchi, Nicolai Marquardt, and Nadia Bianchi-Berthouze. 2016. RealPen: Providing Realism in Handwriting Tasks on Touch Surfaces Using Auditory-Tactile Feedback. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (Tokyo, Japan) *(UIST '16)*. Association for Computing Machinery, New York, NY, USA, 195–205. https://doi.org/10.1145/2984511.2984550

[14] Cornelsen. 2021. *Jo-Jo Fibel.* https://www.cornelsen.de/reihen/jo-jo-fibel-360000360000 (Accessed on 08/25/2021).

[15] Jérémy Danna and Jean-Luc Velay. 2015. Basic and supplementary sensory feedback in handwriting. *Frontiers in Psychology* 6 (2015), 169. https://doi.org/10.3389/fpsyg.2015.00169

[16] Tobias Drey, Patrick Albus, Simon der Kinderen, Maximilian Milo, Thilo Segschneider, Linda Chanzab, Michael Rietzler, Tina Seufert, and Enrico Rukzio. 2022. Towards Collaborative Learning in Virtual Reality: A Comparison of Co-Located Symmetric and Asymmetric Pair-Learning. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 610, 19 pages. https://doi.org/10.1145/3491102.3517641

[17] Tobias Drey, Fabian Fischbach, Pascal Jansen, Julian Frommel, Michael Rietzler, and Enrico Rukzio. 2021. To Be or Not to Be Stuck, or Is It a Continuum? A Systematic Literature Review on the Concept of Being Stuck in Games. *Proc. ACM Hum.-Comput. Interact.* 5, CHI PLAY, Article 229 (Oct. 2021), 35 pages. https://doi.org/10.1145/3474656

[18] Tobias Drey, Jan Gugenheimer, Julian Karlbauer, Maximilian Milo, and Enrico Rukzio. 2020. VRSketchIn: Exploring the Design Space of Pen and Tablet Interaction for 3D Sketching in Virtual Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376628

[19] Tobias Drey, Pascal Jansen, Fabian Fischbach, Julian Frommel, and Enrico Rukzio. 2020. Towards Progress Assessment for Adaptive Hints in Educational Virtual Reality Games. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3334480.3382789

[20] Tobias Drey, Michael Rietzler, and Enrico Rukzio. 2021. Questionnaires and Qualitative Feedback Methods to Measure User Experience in Mixed Reality. arXiv:2104.06221 [cs.HC]

[21] EducationalAppStore. 2021. *10 Best Writing Apps for Kids - Download Now.* https://www.educationalappstore.com/best-apps/writing-apps-for-kids (Accessed on 08/12/2021).

[22] elago. 2022. *Classic Pencil Case for Apple Pencil 2nd Generation.* https://www.elago.com/products/apple-pencil-2nd-generation-cover-yellow-9f4lc?_pos=4&_sid=462ed4284&_ss=r (Accessed on 07/15/2022).

[23] Per Engzell, Arun Frey, and Mark D. Verhagen. 2021. Learning loss due to school closures during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences* 118, 17 (2021). https://doi.org/10.1073/pnas.2022376118 arXiv:https://www.pnas.org/content/118/17/e2022376118.full.pdf

[24] Epic Creations, Inc. 2021. *Epic - Kids' Books & Reading.* https://apps.apple.com/us/app/epic-kids-books-reading/id719219382 (Accessed on 08/12/2021).

[25] Andy Field. 2013. *Discovering statistics using IBM SPSS statistics.* sage.

[26] Rosie Flewitt, David Messer, and Natalia Kucirkova. 2015. New directions for early literacy in a digital age: The iPad. *Journal of Early Childhood Literacy* 15, 3 (2015), 289–310. https://doi.org/10.1177/1468798414533560 arXiv:https://doi.org/10.1177/1468798414533560

[27] Sabrina Gerth, Annegret Klassert, Thomas Dolk, Michael Fliesser, Martin H. Fischer, Guido Nottbusch, and Julia Festman. 2016. Is Handwriting Performance Affected by the Writing Surface? Comparing Preschoolers', Second Graders', and Adults' Writing Performance on a Tablet vs. Paper. *Frontiers in Psychology* 7 (2016), 1308. https://doi.org/10.3389/fpsyg.2016.01308

[28] Paperlike GmbH. 2021. *Paperlike.* https://paperlike.com/ (Accessed on 08/26/2021).

[29] Grammaropolis LLC. 2021. *Grammaropolis-Complete Edition.* https://apps.apple.com/us/app/grammaropolis-complete-edition/id580476057 (Accessed on 08/12/2021).

[30] Jessica Guilbert, Denis Alamargot, and Marie-France Morin. 2019. Handwriting on a tablet screen: Role of visual and proprioceptive feedback in the control of movement by children and adults. *Human Movement Science* 65 (2019), 30–41. https://doi.org/10.1016/j.humov.2018.09.001 Special issue: Articles on graphonomics.

[31] John Hardy. 2012. Experiences: A Year in the Life of an Interactive Desk. In *Proceedings of the Designing Interactive Systems Conference* (Newcastle Upon Tyne, United Kingdom) *(DIS '12)*. Association for Computing Machinery, New York, NY, USA, 679–688. https://doi.org/10.1145/2317956.2318058

[32] Maxwell Holland. 2021. *The 7 Best Apps That Teach Kids How to Read and Write.* https://www.makeuseof.com/apps-teach-kids-read-write/ (Accessed on 08/12/2021).

[33] HPHLP. 2021. *Grammar Pop.* https://apps.apple.com/us/app/grammar-pop/id720189154 (Accessed on 08/12/2021).

[34] Samsul Huda, Nobuo Funabiki, Minoru Kuribayashi, and Wen-Chung Kao. 2019. A proposal of calligraphy learning assistant system with letter portion practice function using projection mapping. *International Journal of Web Information Systems* (2019).

[35] Samsul Huda, Nobuo Funabiki, Minoru Kuribayashi, and Wen-Chung Kao. 2020. A Calligraphy Learning Assistant System with Letter Portion Practice Function Using Projection Mapping. In *2020 IEEE International Conference on Consumer Electronics (ICCE)*. 1–2. https://doi.org/10.1109/ICCE46568.2020.9043042

[36] Apple Inc. 2021. *Apple Pencil.* https://www.apple.com/apple-pencil/ (Accessed on 08/26/2021).

[37] Apple Inc. 2021. *iPad.* https://www.apple.com/ipad/ (Accessed on 08/26/2021).

[38] Sasa Junuzovic, Kori Inkpen, Tom Blank, and Anoop Gupta. 2012. *IllumiShare: Sharing Any Surface.* Association for Computing Machinery, New York, NY, USA, 1919–1928. https://doi.org/10.1145/2207676.2208333

[39] Michelle Kaffenberger. 2021. Modelling the long-run learning impact of the Covid-19 learning shock: Actions to (more than) mitigate loss. *International Journal of Educational Development* 81 (2021), 102326. https://doi.org/10.1016/j.ijedudev.2020.102326

[40] Shaun K. Kane, Daniel Avrahami, Jacob O. Wobbrock, Beverly Harrison, Adam D. Rea, Matthai Philipose, and Anthony LaMarca. 2009. Bonfire: A Nomadic System for Hybrid Laptop-Tabletop Interaction. In *Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology* (Victoria, BC, Canada) *(UIST '09)*. Association for Computing Machinery, New York, NY, USA, 129–138. https://doi.org/10.1145/1622176.1622202

[41] Markus Kiefer, Stefanie Schuler, Carmen Mayerflewitt, Natalie M Trumpp, Katrin Hille, and Steffi Sachse. 2015. Handwriting or typewriting? The influence of pen-or keyboard-based writing training on reading and writing performance in preschool children. *Advances in cognitive psychology* 11, 4 (2015), 136.

[42] David E Kieras and Susan Bovair. 1986. The acquisition of procedures from text: A production-system analysis of transfer of training. *Journal of Memory and Language* 25, 5 (1986), 507–524. https://doi.org/10.1016/0749-596X(86)90008-2

[43] Jeongyun Kim, Jonghoon Seo, and Tack-Don Han. 2014. AR Lamp: Interactions on Projection-Based Augmented Reality for Interactive Learning. In *Proceedings of the 19th International Conference on Intelligent User Interfaces* (Haifa, Israel) *(IUI '14)*. Association for Computing Machinery, New York, NY, USA, 353–358. https://doi.org/10.1145/2557500.2557505

[44] Lenovo. 2021. *Active Pen 2 by Lenovo.* https://www.lenovo.com/us/en/p/accessories-and-software/stylus-pens-and-supplies/stylus-pens-&-supplies_pens/4x81b07782 (Accessed on 08/29/2021).

[45] Letterschool Enabling Learning B.V. 2021. *LetterSchool - Learn to Write!* https://apps.apple.com/us/app/letterschool-learn-to-write/id481067676 (Accessed on 08/12/2021).

[46] Zhen Li, Michelle Annett, Ken Hinckley, Karan Singh, and Daniel Wigdor. 2019. *HoloDoc: Enabling Mixed Reality Workspaces That Harness Physical and Digital Content.* Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300917

[47] Natan Linder and Pattie Maes. 2010. LuminAR: Portable Robotic Augmented Reality Interface Design and Prototype. In *Adjunct Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology* (New York, New York, USA) *(UIST '10)*. Association for Computing Machinery, New York, NY, USA, 395–396. https://doi.org/10.1145/1866218.1866237

[48] Logitech. 2021. *Logitech VR Ink Pilot Edition - VR Stylus for Creativity, Control & Precision.* https://www.logitech.com/en-roeu/promo/vr-ink.html (Accessed on 08/13/2021).

[49] Emilie Loup-Escande, Rémy Frenoy, Guillaume Poplimont, Indira Thouvenin, Olivier Gapenne, and Olga Megalakaki. 2017. Contributions of mixed reality in a calligraphy learning task: Effects of supplementary visual feedback and expertise on cognitive load, user experience and gestural performance. *Computers in Human Behavior* 75 (2017), 42–49. https://doi.org/10.1016/j.chb.2017.05.006

[50] magic leap. 2021. *Safety Guide.* https://assets.ctfassets.net/bl73eiperqoo/1YtAyxfY1ef8pUE7suWH1K/ecf7610f05c2b1dee2abf6eda80f9d62/All_languages_-_Safety_Guide__ML1___Final_.pdf (Accessed on 08/13/2021).

[51] Christian Marquardt, Marianela Diaz Meyer, Manuela Schneider, and René Hilgemann. 2016. Learning handwriting at school – A teachers' survey on actual problems and future options. *Trends in Neuroscience and Education* 5, 3 (2016), 82–89. https://doi.org/10.1016/j.tine.2016.07.001

[52] M. Martens, G. C. Rinnert, and C. Andersen. 2018. Child-Centered Design: Developing an Inclusive Letter Writing App. *Frontiers in Psychology* 9 (2018), 2277. https://doi.org/10.3389/fpsyg.2018.02277

[53] Carmen Mayer, Stefanie Wallner, Nora Budde-Spengler, Sabrina Braunert, Petra A. Arndt, and Markus Kiefer. 2020. Literacy Training of Kindergarten Children With Pencil, Keyboard or Tablet Stylus: The Influence of the Writing Tool on Reading and Writing Performance at the Letter and Word Level. *Frontiers in Psychology* 10 (2020), 3054. https://doi.org/10.3389/fpsyg.2019.03054

[54] Madhura Mhatre, Siddhi Choudhari, and Nishtha Chourasia. 2015. Microsoft pixelsense—Interactive surface computing platform for Samsung SUR40. *Int. J. Res. Eng. Technol.* 4, 11 (2015), 108–110.

[55] Microsoft. 2021. *Product safety warnings and instructions.* https://support.microsoft.com/en-us/topic/product-safety-warnings-and-instructions-726eab87-f471-4ad8-48e5-9c25f68927ba (Accessed on 08/13/2021).

[56] Masashi Narita and Takafumi Matsumaru. 2015. Calligraphy-stroke learning support system using projection. In *2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN).* 640–645. https://doi.org/10.1109/ROMAN.2015.7333576

[57] Michelle M. Neumann. 2018. Using tablets and apps to enhance emergent literacy skills in young children. *Early Childhood Research Quarterly* 42 (2018), 239–246. https://doi.org/10.1016/j.ecresq.2017.10.006

[58] M. M. Patchan and C. S. Puranik. 2016. Using tablet computers to teach preschool children to write letters: Exploring the impact of extrinsic and intrinsic feedback. *Computers & Education* 102 (2016), 128 – 137. https://doi.org/10.1016/j.compedu.2016.07.007

[59] Iulian Radu and Bertrand Schneider. 2019. What Can We Learn from Augmented Reality (AR)? Benefits and Drawbacks of AR for Inquiry-Based Learning of Physics. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19).* Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300774

[60] Hugo Romat, Andreas Fender, Manuel Meier, and Christian Holz. 2021. Flashpen: A High-Fidelity and High-Precision Multi-Surface Pen for Virtual Reality. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR).* 306–315. https://doi.org/10.1109/VR50410.2021.00053

[61] T. Shichinohe, T. Yamabe, T. Iwata, and T. Nakajima. 2011. Augmented Calligraphy: Experimental Feedback Design for Writing Skill Development. In *Proceedings of the Fifth International Conference on Tangible, Embedded, and Embodied Interaction* (Funchal, Portugal) *(TEI '11).* ACM, New York, NY, USA, 301–302. https://doi.org/10.1145/1935701.1935769

[62] solocode GmbH. 2022. *ANTON - Learn for free with lessons in Language Arts, Math, German, Science, Music, Kindergarten to Grade 6.* https://anton.app/en_us/ (Accessed on 04/16/2022).

[63] Rachanart Soontornvorn, Hiroyuki Fujioka, and Takeshi Shimoto. 2016. A development of TANZAKU calligraphy training system using Augmented Reality and dynamic font. In *2016 IEEE Region 10 Conference (TENCON).* 3720–3723. https://doi.org/10.1109/TENCON.2016.7848754

[64] Rachanart Soontornvorn, Sarayoot Pongkarn, and Hiroyuki Fujioka. 2015. A Development of AR Calligraphy Skill Training System Using Dynamic Font. In *2015 10th International Conference on Broadband and Wireless Computing, Communication and Applications (BWCCA).* 555–558. https://doi.org/10.1109/BWCCA.2015.6

[65] Jill Staake. 2021. *The Best Writing Apps for at Home and in the Classroom.* https://www.weareteachers.com/writing-apps-for-kids/ (Accessed on 08/12/2021).

[66] Niels A Taatgen. 2013. The nature and transfer of cognitive skills. *Psychological review* 120, 3 (2013), 439.

[67] Unity Technologies. 2021. *Unity Homepage.* https://unity.com (Accessed on 08/23/2021).

[68] John Underkoffler, Brygg Ullmer, and Hiroshi Ishii. 1999. Emancipated Pixels: Real-World Graphics in the Luminous Room. In *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '99).* ACM Press/Addison-Wesley Publishing Co., USA, 385–392. https://doi.org/10.1145/311535.311593

[69] Vuforia. 2021. *Vuforia Developer Portal.* https://developer.vuforia.com/ (Accessed on 08/24/2021).

[70] Wacom. 2021. *Wacom stylus.* https://www.wacom.com/en-us/products/stylus (Accessed on 08/29/2021).

[71] John Wann and Ian Nimmo-Smith. 1991. The control of pen pressure in handwriting: A subtle point. *Human Movement Science* 10, 2 (1991), 223–246. https://doi.org/10.1016/0167-9457(91)90005-I

[72] Pierre Wellner. 1993. Interacting with Paper on the DigitalDesk. *Commun. ACM* 36, 7 (July 1993), 87–96. https://doi.org/10.1145/159544.159630

[73] Roy Winata. 2021. *iWriteWords (Handwriting Game).* https://apps.apple.com/us/app/iwritewords-handwriting-game/id307025309 (Accessed on 08/12/2021).

[74] Christian Winkler, Julian Seifert, Christian Reinartz, Pascal Krahmer, and Enrico Rukzio. 2013. Penbook: Bringing Pen+Paper Interaction to a Tablet Device to Facilitate Paper-based Workflows in the Hospital Domain. In *Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces* (St. Andrews, Scotland, United Kingdom) *(ITS '13)*. ACM, New York, NY, USA, 283–286. https://doi.org/10.1145/2512349.2512797

[75] Jacob O. Wobbrock and Julie A. Kientz. 2016. Research Contributions in Human-computer Interaction. *Interactions* 23, 3 (April 2016), 38–44. https://doi.org/10.1145/2907069

[76] Yingfei Wu, Xiaoyan Lu, Dibin Zhou, and Yizhou Cai. 2013. Virtual Calligraphic Learning and Writing Evaluation. In *2013 Sixth International Symposium on Computational Intelligence and Design*, Vol. 2. 108–111. https://doi.org/10.1109/ISCID.2013.141

[77] Yingfei Wu, Zhenming Yuan, Dibin Zhou, and Yizhou Cai. 2013. A Mobile Chinese Calligraphic Training System Using Virtual Reality Technology. *AASRI Procedia* 5 (2013), 200–208. https://doi.org/10.1016/j.aasri.2013.10.079 2013 AASRI Conference on Parallel and Distributed Computing and Systems.

[78] Robert Xiao, Scott Hudson, and Chris Harrison. 2017. Supporting Responsive Cohabitation Between Virtual Interfaces and Physical Objects on Everyday Surfaces. *Proc. ACM Hum.-Comput. Interact.* 1, EICS, Article 12 (June 2017), 17 pages. https://doi.org/10.1145/3095814

[79] Ahmet Yamaç, Ergün Öztürk, and Neşet Mutlu. 2020. Effect of digital writing instruction with tablets on primary school students' writing performance and writing knowledge. *Computers & Education* 157 (2020), 103981. https://doi.org/10.1016/j.compedu.2020.103981