Augmented Games – Exploring Design Opportunities in AR Settings With Children

Sascha Oberhuber¹, Tina Kothe², Stefan Schneegass³, Florian Alt¹

¹LMU Munich, Ubiquitous Interactive Systems Group, florian.alt@ifi.lmu.de
²LMU Munich, Institute of Art Education, tina.kothe@lrz.uni-muenchen.de
³University of Stuttgart, Visualization and Interactive Systems, stefan.schneegass@vis.uni-stuttgart.de

ABSTRACT
In this paper we investigate how Augmented Reality (AR) technology influences children during creative content generation in playful settings. The work is motivated by the recent spread of AR and the fact that children get in touch with this technology through their smart phones very early on. To understand the consequences, we implemented an app for smart mobile devices that allows children to create treasure hunts using GPS coordinates and marker-based AR functionality. During a qualitative user study, we asked students (n=27) to create traditional (paper + art supplies) and digital (paper + art supplies + AR app) treasure hunts and compared the resulting games, among other metrics, in terms of complexity, length and types of media used. Whereas traditional treasure hunts were linear, centered around locations and delivered information with text only, digital treasure hunts were more complex, focused on visual aspects and frequently integrated storytelling.

INTRODUCTION
Mobile devices have become ubiquitous, and as a consequence enable novel AR applications in daily live. Such applications are not limited to games and navigation [16], but AR applications have been designed for use in, for example, museums [26], advertising [3], or sports [6]. While regular tablet and phone apps have already proven valuable for education [24, 25], AR app also found their way into the education of children [1, 7, 18]. In these cases, AR is used as a novel and interactive representation of content. For example, ARBS (Augmented Reality Books System) enhances traditional textbooks with AR content [11]. This augmentation has shown to evoke positive emotions and more motivation for the discussed topics [10]. Particularly, providing meta information was considered useful via AR [9, 12]. In another project, the creation of simulations in AR has shown to support children in learning by being highly engaging [8].

AR has mainly been used as a visualization technique by teachers and educators. It was found that AR interfaces may support collaborative activities [4] and lead to higher motivation [22]. At the same time, providing children means to create their own AR content received little attention. Prior projects showed how to build tools that allows children to create digital content in general, for example, programming environment for children (‘Scratch’¹) or story authoring tools (‘Fiabot!’ [20]).

To close this gap, we focus on enabling children to create AR content by means of an AR authoring tool. We investigate how children use the tools to create AR content on mobile devices and how this influences their creative process [14] compared to traditional tools. We explore how the children’s ability to grasp and exploit the concepts of AR helps them to create new content and how it impacts on motivation and experience.

¹Scratch Website: https://scratch.mit.edu

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

IDC’17, June 27–30, 2017, Stanford, CA, USA.
Copyright is held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-1-4503-4529-3/17/06 ...$15.00.
http://dx.doi.org/10.1145/3077548.3077561
To investigate the potential and challenges of AR in an educational content creation process, we first implemented 'MakeAR', an authoring tool that allows AR content to be created. We then let children create traditional (paper + art supplies) and digital (paper + art supplies + MakeAR tool) treasure hunts, qualitatively comparing the resulting games with regard to complexity, length, and types of media used.

CONTRIBUTION STATEMENT
The contribution of this work is twofold. First, we report on the design and implementation of the 'MakeAR' tool allowing digital AR treasure hunts to be created. Second, we report on a user study with 27 fourth grade primary schools children. We observe their creative work process to compare their creative work to traditional non-digitally content creation processes.

MAKEAR – CREATING AR TREASURE HUNTS
'MakeAR' is an app we designed to support children in creating and playing digitally enhanced treasure hunts using augmented reality. The digital enhancement happens through marker-based and/or location-based AR content generated by the user within the application. We developed 'MakeAR' as an Android app particularly suited for tablet-sized devices.

Requirements

While there is a selection of tools allowing users to generate AR content by themselves, our specific setting and target audience (i.e., children aged 9–11) comes with a different set of requirements. The tool should provide:

- simple mechanisms that allow children to focus on the content (instead of the program).
- content generation at runtime, including AR overlays, GPS hooks, and visual markers.
- an approach not relying on an active internet connection to allow usage in places with limited connectivity (e.g., in the woods or basement).

Existing solutions like Aurasma do not include these requirements and were not designed for being used by children.

General Workflow

On opening 'MakeAR' provides two buttons for the main functions: create and play. When choosing create, users see empty fields that can be tapped to add a station to the treasure hunt. A simple treasure hunt is shown in Figure 2. A station must be linked by a cordlike line to the next station. These stations can be dragged and dropped to different fields, the lines be cut and reassigned to other stations. Branches and multiple different endings are possible. Each station requires one trigger and at least one piece of AR content. A trigger is either a set of GPS coordinates (current position, stored in the app) or a hand-drawn paper marker (picture with a black, roughly square outline to be recognized as fitting marker). The paper marker must be scanned by the app and is then transfigured into an internal digital marker, used for image recognition. AR content can be added after the trigger (GPS or visual marker) is assigned to a station. Content can be created in three different ways: (1) A photo from the surroundings, objects of interest or clues can be taken by an integrated camera interface. (2) Children can write or hand-draw on paper and digitalize it by taking a photo. (3) Images (text or pictures) can be imported from Android’s file storage. This allows modification of photographs through image editing or integrating pictures from third-party drawing apps.

Playing a previously created treasure hunt is the other main function of 'MakeAR'. Children hand over their tablet with their previously created treasure hunt stored on it. They give a hint to the playing group where to find the first clue – the rest is provided by the app. In addition to whatever information the content provides, stations that use GPS triggers can be located with the help of a compass integrated in the app.

When the children approach a GPS-based clue, a treasure chest is integrated into the live camera view (see Figure 3). The chest’s location is recalculated based on the movement of the tablet, and the actual content is collected by tapping on the chest, which opens it. Marker-based clues work by approaching the physical marker in the real world and tapping the augmented content that is projected on top of it. All clues that got collected throughout a treasure hunt remain as minimized items on the lower half of the screen, thus allowing the players to review them.

MAIN STUDY

Our research is driven by the question how the use of AR influences creative content generation among children. Hence, we compare a traditional, paper-based treasure hunt game with an AR-treasure hunt. We follow an exploratory approach with a strong focus on qualitative observations. We also collected some quantitative data, yet decided to report descriptive statistics only due to the rather small sample size.
Study Design & Participants
The study followed a between subject design where half of the children designed a treasure hunt using traditional media such as pen, paper, and duct tape, whereas the other half was offered both traditional tools and the MakeAR app described above. The study was conducted over the course of two days, with day 1 dedicated to the ‘digital treasure hunts’ using MakeAR, and the traditional treasure hunt taking place on day 2. During each day, only the children taking part in the day’s experiment were present; nobody participated on both days.

The children were recruited among fourth graders from a local elementary school and were aged 9 to 11. In total, 14 children participated in the digital treasure hunt. In this group we had 6 boys and 8 girls, split into four teams of one boys team, one girls team and two mixed teams. There were 13 children recruited for the traditional treasure hunt (3 girls, 10 boys). Here the children split into three boy-teams and one girl team.

Note, that an earlier prototype of ‘MakeAR’ (using GPS localization instead of AR markers) was previously tested among children from the same school with regard to usability [23]. We recruited participants in the digital treasure hunt group from among these children. Thus, participants were already familiar with the app and we reduced a potential novelty bias.

Procedure
The researchers started both days by presenting a short overview about the idea of the project and verified that everyone was familiar with the concept of treasure hunt games. The children’s task was to create a treasure hunt for their classmates in a given timeframe using the tools provided by the researchers. On the first day, the children were additionally introduced to ‘MakeAR’s visual marker functionality with a small example, a three-dimensional apple superimposed on a black and white apple marker. Then, the children split into four teams, to each of which one observer from the team of researchers was assigned. Observers were instructed to intervene as little as possible, in particular regarding content creation and children’s approach to solving the task.

To clarify the task, we provided two short examples of a treasure hunt on both days. These examples differed in terms of complexity. Both followed a short, overarching story about a woman sending the players to a merchant to get a sewing thread to fix her dress. However, while the first example was strictly linear, the second example included different branches and a recall task – the dress was colored and children had to choose between different sewing threads (blue or red) when reaching the merchant. Based on their choice, the last location and the end of the story differed. Every team only completed one example treasure hunt (linear or complex). This allowed us to retroactively analyze whether children would adhere to what they had seen, or come up with something different.

After completing the examples, each team started to create a treasure hunt by themselves (Figure 4). No further instructions were given how the treasure hunt should look like and which elements it should contain. Teams were provided with various types of paper, pencils, colored pens, scissors, duct tape, and further drawing supplies. In addition, the AR group was provided a Samsung Galaxy Note Pro 12.2 tablet with ‘MakeAR’ installed. Once content was created (60–90 minutes), every team played another team’s treasure hunt.

After that, children were asked to reflect on their experiences, issues, and ideas for improvement during an open group discussion. Children were also handed out questionnaires (Scales of Enjoyment [19]) that asked how much they enjoyed creating their own game, playing the others’ game, and the overall experience. Open questions inquired about how they would describe their own game and whether they had further ideas to improve their game or the others’ game.

RESULTS
Overall, 7 of 8 teams (traditional and digital) managed to create a complete treasure hunt and to play another team’s treasure hunt. One digital team was so immersed in the design and story of their artifacts that they forgot to properly link their stations logically, which was fixed post-hoc using paper-based hints attached to the marker images. This team also managed to play another team’s treasure hunt.

Examples
We provide some examples of hunts created by both groups.

Traditional Group
One of the traditional teams created a game that sent players across the school campus. In this treasure hunt, artifacts were pieces of paper with nothing but directional text instructions like "look for the big tree" or "go to place XY" (Figure 6). Another traditional treasure hunt team created a game with less stations, but put more emphasis on the individual artifacts. Most instructions were encoded in riddles (simple cryptography such as mirrored text or letter substitution). In contrast to the previous group they intentionally hid their artifacts well.
Figure 5. Graph-example of a simple digital treasure hunt: The games were transformed into graphs, each node representing one station. Nodes consist of one or more artifact. Each artifact is described by the artifact type (T), its content (C) and the medium (M) used to create it. This allowed the level of complexity and type of solution of the treasure hunt to be assessed.

Figure 6. These exemplary traditional hints, send the players to "where the butterflies are" and "to the dressers". They are weakly linked hints consisting of nothing but text.

Figure 7. A highly sophisticated artifact, containing a modified photo and a textual hint wrapped up in a story.

**AR Group**

One AR team built a treasure hunt that contained most types of artifacts: marker-based stations, GPS, and physical riddles. Two artifacts were based on photographs the children took and on which they afterwards placed text and color with a drawing app (Figure 7). The stations were tied together by a short story about a stolen pacifier that had to be recaptured.

The second AR team also came up with a story: here a thief needed to get medicine for his ill horse and later bought supplies at the black market. The game was exclusively built using visual markers that led to dialogs among the characters, played by the members of the team. As previously mentioned, these students were so immersed in realizing their story that they forgot to leave hints where to find the next station.

**Game Complexity Analysis – Quantitative Findings**

During the study and the following analysis we observed some differences in the children’s approach to AR treasure hunts as opposed to traditional hunts. To understand these differences on a quantitative level, we analysed the complexity of the developed games. The criteria were based on indications for creativity, such as preferring complexity over linear solutions, story over simple instructions, and sophisticated images over text-based one-liners [14, 20].

To understand the complexity, the level of sophistication, and story-driven content we analyzed the graph generated through each treasure hunt (see Figure 5 for an example).

We investigated the content of the treasure hunt such as the number of stations and artifacts (e.g., physical, digital, or GPS coordinate) as well as what kind of content was used (text and/or images, riddles, etc.). We also looked into how children created these pieces of content and which media (pen, paper, photo) they used. Note that in the AR version, content could also be created with pen and paper and then be integrated by taking a picture with MakeAR. This option was the dominant way of usage. We also investigated the connectivity between the stations (grouped into weak, strong, and recall link). The results are summarized in Table 1.

**Number of stations and artifacts.** The traditional group created 35 stations in total, each consisting of exactly one artifact: a hint to the next station ($M=8.8, SD=1.7$). The AR group created 21 stations ($M=5.0, SD=1.2$), consisting of 38 artifacts ($M=9.5, SD=1.3$). Of these artifacts, 14 were markers, 2 were GPS coordinates and 5 were physical objects, such as small drawings pinned to walls and texts reading: "winner!!!" or "Game over :-( Try again!".

**Content.** All but one artifact (total of 35) created by the traditional group focused on text. 11 of them contained riddles. 2 artifacts featured images, one of which contained only an image and no text.

In the AR group, 22 artifacts (total of 38) contained text (including one riddle) and 19 contained images. In total 3 of the 4 teams embedded all of their artifacts into a story. Note that these numbers do not add up to the previously stated 38 artifacts because some artifacts featured both text and imagery.

**Connectivity.** Finally, we looked at how the content of the stations was connected. We differentiate between "weak" and "strong" links, and define a weak link to be a simple directional instruction ("The next hint is at the large tree."). Strong links, on the other hand embed directional information into other elements, for example an overarching story ("You can hear someone scream. Its coming from the classroom 4A!"). Furthermore, links may reference previous artifacts.
In the traditional group, stations were connected through weak links exclusively. In contrast, the AR group strongly linked their artifacts (10) through stories. Overall, we found that in the traditional group, children utilized 31 weak links, whereas in the AR group, children built 4 weak, 10 strong, and 1 recall link (players had to recall an artifact from a former station).

**Qualitative Findings**

The qualitative analysis was based on a review of the treasure hunts themselves (game design, group interaction), the created artifacts (digital and physical), and observer notes. Additionally, we collected data via questionnaire and recorded videos during the game creation phase and final group discussions.

**Game Design**

In both evaluated groups, it was possible to design artifacts with text and images. We found that with two exceptions, members of the traditional group did not include imagery into their artifacts but almost exclusively used text (Figure 4–a and Figure 6). During the digital treasure hunt, half of the artifacts contained imagery (Figure 4–d and Figure 7). Children often took multiple attempts to draw markers for the sole reason of making them look good (Figure 8), as opposed to aiming for mere functionality. With the exception of a single drawing (a treasure chest on a “goal”-type artifact), traditional groups did not include any elements that would qualify as visual design.

**Quantitative Findings**

We found that the AR group invested about 270 minutes in total to create 38 artifacts, amounting to 0.6 artifacts created per child per hour. In the traditional group, children created on average 0.9 artifacts per hour, hence being about 50% faster in content creation. Members of the AR group spent the additional time in the sophisticated design of the artifacts as well as in the structure of the games themselves. This is also reflected by the fact that one of the AR teams first sketched the story on a blackboard before implementing the treasure hunt (Figure 4–d). Also text-based artifacts originating in the AR treasure hunt group tended to be longer (more words and letters) and hold more information (Figure 7).

Even though every AR marker was treated as image by our system, text as content is similarly detected and makes a valid marker as well. Thus, the children could place text on an AR marker. These text-based markers were counted as ‘text-based artifacts’ for the purpose of this study. It was also possible to create an AR treasure hunt by using GPS coordinates instead of visual markers and therefore omitting imagery completely. However, the students did not create such AR treasure hunts.

Table 1. Quantitative results of the study. We compared the groups with regard to number of stations, artifacts and connections they created.

<table>
<thead>
<tr>
<th></th>
<th>Traditional Treasure Hunt</th>
<th>AR Treasure Hunt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of stations</td>
<td>35</td>
<td>21</td>
</tr>
<tr>
<td>Artifacts</td>
<td>35</td>
<td>28</td>
</tr>
<tr>
<td>Text</td>
<td>34 (97%)</td>
<td>22 (58%)</td>
</tr>
<tr>
<td>Image</td>
<td>2 (6%)</td>
<td>19 (50%)</td>
</tr>
<tr>
<td>Riddle</td>
<td>11 (32%)</td>
<td>1 (3%)</td>
</tr>
<tr>
<td>Connections</td>
<td></td>
<td></td>
</tr>
<tr>
<td>weak</td>
<td>31</td>
<td>4</td>
</tr>
<tr>
<td>strong</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>recall</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Artifacts created per h</td>
<td>0.9</td>
<td>0.6</td>
</tr>
</tbody>
</table>

One of the four traditional teams used a lot of textual riddles compared to the short and concise directional statements that dominated the other traditional groups’ games.

**Storytelling vs. Movement & Location**

Three out of four AR teams focused on telling stories through their artifacts. This was reflected both through the artifact design as well as by the way in which the stations and also the artifacts within the stations were connected. The AR teams created stronger links: directional instructions were usually embedded into the story (Figure 7). In contrast, none of the traditional treasure hunt teams made an attempt to tell a story. They rather focused on the sequence of locations and created artifacts serving as direct pointers between the locations.

It is worth noting that one of the four AR teams chose to build a treasure hunt with little logical links, where the creation process was dominated by coming up with places to deposit artifacts. This shows that the children of the AR teams were aware of the possibility to create linear text-based treasure hunts, but in 75% chose to focus on storytelling.

**Questionnaires**

The children rated questions on 4-Point Likert items (Smiley-o-meter–1=totally agree; 4=totally disagree). When asked whether they enjoyed playing the other group’s treasure hunt, the average rating for the traditional group was 2, whereas it was 1.8 on average for the AR group. There was a slight preference among children to create their own treasure hunts using AR (M=1.15) versus creating it traditionally (M=1.23).

**Suggestions for Features and Improvement**

In the open group discussions, the children suggested various features to improve ‘MakeAR’. This included video function, speech recording, support for different levels of difficulty, time-based challenges, a VR simulation to play indoors in case of rainy or snowy weather, and the ability to stick artifacts to real objects out of reach, such as clouds or treetops.

When asked for improvements for the traditional concept, the children focused solely on error prevention. While playing, many children came across artifacts that didn’t belong to the treasure hunt they were playing, got mixed up with other treasure hunts and thus ended up confused. To avoid this in the future, they suggested using text or colour to mark artifacts. Again, this shows that children were willing to instrumentalize properties of the clues that could otherwise be used for visual design to improve playability.
DISCUSSION
Our main focus was to observe the content creation process. The results suggest a slight preference of the AR groups to compose story-driven, image-based and complex treasure hunts. These are indicators for creative processes such as being able to do complex thinking, to go into varied directions and prefer unusual solutions over straightforward results [14]. This section highlights our conclusions drawn from our study.

The game complexity analysis shows that the AR treasure hunt group preferred a complex setting with many variables over strictly linear solutions. The game design findings reveal that the children focus on visual and aesthetical solutions instead of on plain and straightforward designs. This implies that marker and content design do qualify as what prior work describes as ‘aesthetic experience’, i.e. children focused on visual design, way beyond the ‘utilitarian breakpoint’ of getting working artifacts done [5, 17].

The questionnaire reveals a slight preference for creating over playing games. There can be multiple explanations (from technical issues to other children’s games being not as satisfying as imagined) but one reason may be the new and pleasing experience of creating something digital themselves as opposed to using existing apps and media. This conclusion would be in line of thinking with media literacy concepts such as ‘active media education’ [13]. The group discussions displayed a wide imagination of the AR group towards additional features, other ideas, solutions, and room for improvement. This kind of thinking indicates a creative mind and could therefore be an implication for creative problem solving strategies [2, 14, 17].

The results show that using pen and paper seems to motivate children to quickly execute ideas and concepts (in this case, to create many stations). At the same time, less care was exercised regarding the design of each station or the overall game design. This approach can be useful in case prototyping with quick results are needed in early stages of a design process.

Children also employed a structured working process using the MakeAR application (sketching the story, thinking about and building artifacts, linking artifacts together, and blending them in the story). In the traditional setting, children switched back and forth between discussing possible locations and creating artifact which were then rather weakly linked. This could be an indication that AR settings support structured working.

Our results show that creating AR content influences the working atmosphere in a similar manner as when presenting content [10]. It positively influenced children’s motivation, evokes positive emotions and is highly engaging. This is reflected by the high number and quality of the created artifacts and the positive attitude children showed during the study.

Finally, the high level of detail and the strong focus on stories hints at AR triggering self-initiated learning. Without the observers’ or educators’ input, the students used the method of a structured working process, attempted and discarded ideas in consensus with their team and managed to create games in the given time span by themselves. This type of thinking and acting benefits the children’s development and learning abilities in an educational setting [2, 14, 21].

LIMITATIONS AND FUTURE WORK
We acknowledge several limitations of our work. Our study was conducted with a limited number of children and had a strong exploratory focus. As a result we only report descriptive statistics. We see our findings as first implications that point to interesting directions for future work. Follow-up work could pick up specific aspects and explore them in-depth using a thorough quantitative approach. Such an approach could take into account potential gender differences.

A second limitation is that some differences we observed between the two groups may not be a result of using AR only. For example, the ‘MakeAR’ software provides a scaffold for creating the treasure hunt, which may have supported the more structured process. Furthermore, prior work in art education indicates that using digital surfaces (such as tablets or PCs) invokes different types of visual interactions (for example, drawing, painting, staging) compared to working on the same task using paper and pens [15].

Interestingly, the AR groups mostly chose to create story-driven treasure hunts, even though this was not the given task. Furthermore, the ‘MakeAR’ app seems to foster image-based content and storylines. This should be investigated more closely, perhaps in comparison to other AR-creation tools.

Further scenarios (away from treasure hunts) need to be evaluated to obtain knowledge about creative processes through AR in different educational settings.

CONCLUSION
We explored how children handle a creative, complex, yet playful task that involves AR technology and compare it to a traditional pen and paper approach. We did so by, firstly, implementing an AR application for mobile devices that offers opportunities for children to enhance treasure hunts with AR elements. Secondly, we conducted a study involving school children aged 9-11 and asked them to create treasure hunts.

Findings from our qualitative analysis suggest that AR technology may indeed be used to complement educational settings that focus on qualitative outcomes by supporting creativity, self-initiated learning, and structured working processes of the children while at the same time creating a motivating and positive environment. We point to interesting aspects that should be subject to future investigations.

ACKNOWLEDGEMENTS
We thank the 23 children who participated in our user study for their constructive feedback, very reasonable criticism, and enthusiastic suggestions. We also thank Christian Römmelt from the Gänseliesel primary school as well as Grudrun Klinker, David Plecher and Thomas Stiefelbauer of the Augmented Reality Group of TUM for their support.

REFERENCES


17. G. Peez. 2015. *Art Education in Germany*. Waxmann Verlag GmbH.


