DESIGN-BASED RESEARCH FOR THE DEVELOPMENT OF AUGMENTED REALITY APPLICATIONS: THE CASE OF VIRTUAL PLAYGROUND

Pesquisa baseada em design para o desenvolvimento de aplicações de realidade aumentada: o caso do virtual playground

Manoela Milena Oliveira da Silva  
Doutora em Ciência da Computação  
Universidade Federal de Pernambuco, Voxar Labs – Pernambuco – Brasil  
manoela.milena@gmail.com  
http://orcid.org/0000-0001-5401-6540

João Barbosa Moizes  
Graduando em Design  
Universidade Federal de Pernambuco, Voxar Labs – Pernambuco – Brasil  
jhm5@cin.ufpe.br  
http://orcid.org/0000-0003-4605-1746

Patrícia Smith Cavalcante  
Doutora em Educação  
Universidade Federal de Pernambuco – Pernambuco – Brasil  
patricia3smith@gmail.com  
http://orcid.org/0000-0003-3570-684X

Verônica Teichrieb  
Doutora em Ciência da Computação  
Universidade Federal de Pernambuco, Voxar Labs – Pernambuco – Brasil  
vt@cin.ufpe.br  
http://orcid.org/0000-0003-4685-3634

Resumo

Os aplicativos de Realidade Aumentada para educação mudam a experiência de sala de aula para professores e alunos. No entanto, o uso dessas ferramentas ainda não é generalizado. As aplicações atuais ainda possuem algumas limitações, como: interfaces difíceis de usar, falta de liberdade para criar conteúdo, falta de suporte para uma experiência compartilhada entre outros. Este trabalho apresenta a atualização do Virtual Playground, um aplicativo criado de forma colaborativa com professores de idiomas, designers e engenheiros; e Virtual Playground Creator,
A ferramenta de autoria para esse aplicativo. Este processo foi desenvolvido seguindo a metodologia de pesquisa baseada em design em dois ciclos de ideação, implementação e teste. Apresentamos as atualizações e resultados referentes à segunda rodada de testes do *Virtual Playground* com seis professores. Os principais objetivos foram avaliar se o *Virtual Playground Creator* permite que os professores criem experiências de RA para usar em sala de aula; e (2) investigar se havia recursos ausentes ou desnecessários no *Virtual Playground Creator*. Os resultados mostraram que a ferramenta permitiu aos professores trabalhar com diversas habilidades e competências. O protótipo do *Virtual Playground Creator* permitiu ao professor criar suas experiências de RA propostas, embora, em alguns casos, fossem necessárias modificações. Os principais problemas foram relacionados à limitada biblioteca 3D no protótipo avaliado e à falta de possibilidades de criação.

**Palavras-chave:** Realidade Aumentada; Pesquisa Baseada em Design; Ensino de Idiomas; Educação; Educação Tecnológica.

**Abstract**

Augmented Reality applications for education change the classroom experience for both teachers and students. However, the use of these tools is not widespread yet. Current applications still have some limitations, such as: hard to use interfaces, lack of freedom to create content, lack of support for a shared experience among others. This work presents the update of Virtual Playground, an application collaboratively created with language teachers, designers and engineers; and Virtual Playground Creator, an authorship tool for this application. This process was developed following Design-based research methodology in two cycles of ideation, implementation and testing. We present the updates and results regarding the second round of six tests of Virtual Playground with teachers. The main goals were to evaluate if Virtual Playground Creator allows teachers to create AR experiences to use in classrooms; and (2) investigate if there were missing or unnecessary features in Virtual Playground Creator. Results have shown that the tool allowed teachers to work with diverse skills and competencies. The Virtual Playground Creator prototype allowed the teacher to create their AR experiences proposed, although, in some cases modifications were needed. The main issues were related to the limited 3D library in the prototype evaluated and lack of creation possibilities.

**Keywords:** Augmented Reality; Design-based Research; Language Teaching; Education; Technological Education.

**INTRODUCTION**

Augmented Reality (AR) has the potential to completely change the learning and teaching experience, it can improve content understanding and enhance student motivation and collaboration. AR has great potential for language teaching, it can help to create context and scenarios where students can practice the language creatively.
(BILLINGHURST et al., 2001). It can seamlessly enhance the real world with virtual content.

Many AR applications are designed for educational purposes, such as Anatomy 4D (ANATOMY4D, 2016) and Wonderscope (WONDERSCOPE, 2020). Also, studies have shown AR's positive impact on learning (RADU, 2014; THEODOROU et al., 2018; RADU et al., 2021). Nonetheless, AR's use is still limited (SILVA et al., 2019). Applications can be time-consuming for the preparation and execution of a lesson plan, the limited content may prevent the teacher from exploring a topic, or the application is only available for platforms with difficult access.

Much of this can be explained because these applications were created with little or no partnership with teachers or students. The classroom environment presents specific challenges. Thus, a tool that intends to help teachers and students must consider these specificities. Thus, this work presents the updates and improvements conducted on an AR application focused on language learning, developed collaboratively with teachers, designers, developers, and researchers. Thus, the contributions of this paper are: (1) the improvement of an educational AR application that involved a multidisciplinary team of an educational researcher, computer science engineers, educators, and designers; (2) an updated version of an educational AR application for language learning (Virtual Playground) and its authorship tool (Virtual Playground Creator); (3) the evaluation of Virtual Playground (VP) with teachers, which validated the application and helped to improve it.

OBJECTIVE

The main goals of this work were to evaluate if Virtual Playground Creator (VP Creator) allows teachers to create AR experiences to use in their classrooms and investigate if there were missing or unnecessary features in Virtual Playground Creator.

METHODOLOGY

The concept's core and the features of the application explained in this work emerged from a Design-Based Research (DBR) inspired process (PEREIRA et al. 2020).
DBR is a process well suited to address the problems and challenges involving technology and education (AMIEL, 2008). It can be used for general insights about the learning process, improving or making new theories, and finding new principles. It combines the advantages of qualitative and quantitative methods while focusing on developing applications that can be built and effectively integrated into social practices, considering its diversity and specific properties (MATTA et al., 2014).

DBR was used and adapted to extract what would be the most important features of an AR application to be used in a classroom environment, as shown in (PEREIRA et al. 2020). This paper describes the process of validation and improvements carried out during the second cycle of the DBR process involving 6 teachers recruited as volunteers from social media communities related to AR for education and inside the University of the authors. The goals of this second round of testing were to: (1) evaluate if VP Creator allows teachers to create AR experiences for VP they planned to use in the classroom and (2) investigate if there were missing or unnecessary features in VP Creator.

This second cycle of research was conducted after the first round of tests described in Pereira et al. (2020), which aimed at validating the concept of the proposed tool. We evolved the Virtual Playground Creator to a functional mobile version aimed to address the feedback from previous evaluations and include design principles discovered up to this point. This version included the modules for the teacher and the student.

Due to the COVID-19 pandemic, these tests needed to be conducted online, enabling access to teachers not necessarily located in our city, thus expanding recruiting possibilities. The tests occurred between April and May 2021.

To test these prototypes, we followed the steps below:

- Teachers read and sign a consent form using Google Forms;
- We presented the research goals, the time required to participate, the equipment needed, and the steps involved in the test;
- Teachers filled out a form about their demographic information and the design principles employed in the tool tested;
- We explained what the VP is and showed parts of the prototype that are available and missing;
- Teachers explored and got familiar with the VP prototype through a live Twitch presentation. They could chat, clarify questions and ask to see the tool from different angles;

- Teachers prepared a didactic sequence using the tool Miro. In which, they had access to cards, illustrated in Figure 01, containing elements to help create a lesson plan (SILVA, 2018). Then, they organized these elements in a Miro canvas, as shown in Figure 02;

- Teachers created what they said they would like students to do with VP using its authoring tool. They downloaded the VP Creator APK and used a screen recorder of their choice to record their screen while using the tool. They sent this recording to the first author. We applied the think-aloud protocol (MARTIN, 2012) while teachers used VP Creator.

- Teachers filled out a Google Form to evaluate the prototypes tested.

Before the test, teachers received a checklist with the equipment needed for the test and a Google Slides presentation summarizing its steps and goals. This protocol was validated and improved through discussions with specialists (i.e: Ph.D. researchers in the areas of technology in education, computer science, and interaction) as well as a pre-test with a designer; and a pilot test with an English teacher (KITCHENHAM, 2002; RUEL, 2016).

Figure 1 - Planning cards used in the second round of testing

Font: Adapted from (SILVA, 2018).
During the test, the data collected from each teacher interviewed consisted of the test transcript, 2 forms, 1 video of the teacher using the VP Creator, and 1 canvas screenshot of the teacher's didactic sequence planning. Teachers are identified by T#, where # is their corresponding number.

Due to the large amount of data, we decided to use the software ATLAS.ti, which belongs to the computer-assisted qualitative data analysis software category, to support the analytical process. This process has been shown to appropriately support qualitative research, easing the coding and visualization, retrieving, and analyzing processes (POCRIFKA, 2019). Our unit of analysis was all the documents of each teacher considered together.

**SOLUTION REFINEMENT**

The first cycle of tests involved 5 language teachers who had previous knowledge about AR and tested both VP and VP Creator to validate their concept. They used the tools and later participated in an interview to share their general thoughts and impressions.
Thinking aloud protocol was also used to reveal their impressions of the interface throughout their experience. These tests are detailed in (PEREIRA et al. 2020).

Based on the results from our first cycle of tests, different issues needed to be fixed to promote a better user experience (PEREIRA et al. 2020). We highlighted the most critical ones and classified them into three different categories, namely:

- Bugs: for example, instability reported with some of the 3D objects;
- System adjustments: for example, the need to create a history of the use;
- Pedagogic adjustments: such as flexibility in defining rules for the activities.

DATA ANALYSIS OF THE SECOND CYCLE

The data was analyzed following the thematic analyses (BRAUN, 2006) and the coding process proposed by Saldaña (2013). This process allows better systematization of the coding process, decreasing subjectivity since the steps and criteria are clearly defined (POCRIFKA, 2019).

First Coding Cycle

Before the first coding cycle, we carried out the pre-coding phase, which involves circling, highlighting, bolding, underlining, or coloring rich or significant participants' quotes or passages that strike the researcher (SALDAÑA, 2013). We used the software ATLAS.ti to highlight those passages. Memos were also written to register the first perceptions of the researcher related to the data at hand.

During the first cycle of coding, Saldaña (2013) offers 24 possibilities of codes. We used grammatical methods. As part of it, we applied attribute coding, which is characterized as a notation, usually at the beginning of a data set rather than embedded within it, of basic descriptive information. We collected participants' demographic information and their profiles (e.g: experience with AR).

We also applied subcoding. We first coded categories related to each research goal and later divided those codes into subcategories, as shown in Table 1. Elemental methods were also used. We applied descriptive coding as part of this method, which summarizes
in a word or short phrase a topic of a passage of qualitative data. These codes came up based on the subcoding created previously.

### Table 1 - Subcoding

<table>
<thead>
<tr>
<th>Code</th>
<th>Nº of Citations</th>
<th>Details About the Code Creation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interface</td>
<td>81</td>
<td>Aspects related to the tool's interface.</td>
</tr>
<tr>
<td>Lesson planning</td>
<td>36</td>
<td>Aspects related to the creation of the didactic sequence.</td>
</tr>
<tr>
<td>Comments about the tool</td>
<td>21</td>
<td>Teacher's comments about the concept and/or use of the tool.</td>
</tr>
<tr>
<td>Teacher profile</td>
<td>28</td>
<td>Aspects related to the teacher's profile.</td>
</tr>
</tbody>
</table>

Font: Elaborated by the author

**Exploratory methods** were also used. As part of this method, we used provisional coding, establishing a predetermined "start list" set of codes before the fieldwork. In our case, we created a code based on an adaptation of the classification of the degree of maturity of the use of the tool in the didactic sequence. Finally, teachers' profiles were classified based on the adopter's categories developed by Rogers (2003) and their time teaching as proposed by Huberman (2000). We also used holistic coding. These codes refer to teachers' comments and characteristics of their profiles, such as their experience and perceptions of technology and content creation. Finally, affective methods were used. We coded compliments made to the tool, which reflect emotions towards it.

**Transition Coding Cycle**

The goal of the transition cycle is not to “take the researcher to the next level” but rather to cycle back to the first coding efforts so the researcher can strategically cycle forward to additional coding and qualitative data analytic methods (SALDAÑA, 2013).
For this cycle, the author offers four possibilities of codes. We chose *eclectic and simultaneous coding*. We refined the codes generated in the first cycle of coding.

**Second Coding Cycle**

Second cycle coding methods are advanced ways of organizing and reanalyzing data coded through first cycle methods (SALDAÑA, 2013). He proposes six types of coding for this cycle. Among the options offered, we chose *pattern coding* to help us achieve the goals of this study, as it is a way of grouping those summaries into a smaller number of sets, themes, or constructs. The final version of the codes is shown in Figure 3.

Figure 3 - Second Cycle Codes. (MAT) for maturity level. (DS) for didactic sequence. (COM) for comments. (TP) for the teacher's profile. (INT) for the interface.

RESULTS OF THE SECOND TEST ITERATION: VALIDATION OF THE ANTICIPATION OF USE
This round of tests was conducted with 6 experienced language teachers. They were familiar with AR and have used it at least for personal use. Some of them participated in the previous steps of the research. The average length of the test was 2h36m34s. The longest test lasted 3h09m47s, and the shortest one lasted 1h38m14s.

**Teacher's profile**

The sample of teachers was balanced regarding gender. 50% of them are male and 50% female. They were aged between 31 and 48 years old. Their average age was 37 years old. As regards their academic background, the majority (4) graduated in language teaching (both English and Portuguese). Whereas 1 graduated in communication and the other in foreign trade. Despite the fact that 2 were not graduated in the area they work in, they are all experienced teachers. Their average teaching time was 14.8 years.

Most participants can be classified in the experimentation stage of their careers, which suggests that they are in a phase characterized by experimentation, motivation, and search for new challenges and/or moments of questioning and reflection on their careers. Only one teacher was classified in the conservatism phase, a phase of either conformism or activism (HUBERMAN, 2000). This particular teacher demonstrated to be active. She is even currently doing a master's degree in education and technology. In general, these teachers seemed very engaged in their profession.

As regards their adoption categories (ROGERS, 2003), 50% were classified as early majority adopters. The other three teachers were classified as innovator, early adopter and late adopter. Regarding their AR knowledge, three used AR primarily as entertainment through games (i.e. Pokémon Go and Snapchat filters) and two used it with students. Two of them have limited knowledge of AR.

**Considerations About Teacher's Sequence (DS)**

To evaluate if VP Creator allows the teachers to create the AR experience for VP they planned, we attempted to answer two questions: (1) are there any relationships between the possibility of applying VP in the didactic sequence and the teacher's
experience with AR? and (2) are there any relationships between the teachers profile and the level of maturity of technology use in their didactic sequence?

As regards question (1), we could not infer a strong relationship between the possibility of applying VP in the didactic sequence and the teacher's experience with AR. Nevertheless, T2 and T6, who used AR in their lessons, previously proposed activities that expanded the current possibilities offered by the tool. Thus, they would need to modify their original planning to apply it. On the other hand, T1 and T4, who reported limited knowledge of AR, did not feel completely comfortable with it at first.

For instance, T4 opted for a more traditional approach named PPP (Presentation, Practice, Production) as illustrated in his speech:

(...) As I mentioned before, I'm thinking about a traditional class in a sense that it doesn't have a lot of things different from what we're used to seeing in classes so that it stays well within control and that it can be done well. It's within what the teacher is used to doing. Then we have a new element, which is generally the use of the tool, right? when the students will use the virtual playground to create their stories. If we were to think about an English class. This lesson would follow a very traditional approach: PPP. Presentation, practice, production. And then, at the time of production, the students would use the virtual playground. And then, they will build their little story. The teacher would use it beforehand to demonstrate, right? To insert them in a story there as an example. And then each student would be invited to use it.

T1, when asked if she had any questions or concerns regarding the use of VP and/or VP Creator, replied the following: "I don't know if I would be able to give the necessary instructions to the students today if they have many questions."

It is worth noting that although T1 and T4 reported limited AR experience, they have plenty of experience using technology for education. T1 is doing a master's degree on that topic, and T4 has been researching it throughout his academic career. This suggests that even though teachers are familiar with other types of technology, this knowledge may not automatically transfer to AR.

Teachers who used AR for entertainment evaluated that they could use VP in their classrooms. T5 mentioned the need to present written information to students during their production time. T6 highlighted the need to adapt the themes chosen for her lessons.
The code comments "COM", shown in Figure 3, registered the teacher's interest in applying the tool in their classrooms. T1, T2, T5, and T6 spontaneously demonstrated their interest in applying it in their classrooms. Although T2 worked as a technology consultant in the school, she was interested in applying it as an invited teacher.

However, the limitations of the 3D library, as expected, were mentioned by all the teachers. T3 highlighted the need to create his own 3D objects. Nevertheless, when asked if he had already created 3D content, he mentioned the following:

Well, create from scratch with an application, with games, no because I don't have the tools nor the institutions I work with have these resources. But it has already been suggested to work with VR at the school I teach, but in fact, we were never able to produce anything. Nor use the games available there.

This quote suggests that the teacher needs more input from the school to both use and create AR and VR content. This shows that although teachers usually desire to create content, this is a difficult task in real life, which might need extra support from stakeholders.

Regardless of their previous use of AR, teachers did not present difficulties in elaborating the didactic sequence. Their rationale for creating the experiences were their previous teaching experiences (T4 and T5), lessons they would be teaching in the next few days (T1, T3, and T6), and their impressions after the observation of the tool (T2).

Question (2) refers to possible relationships between the teacher profile and the level of maturity of technology use in the didactic sequence. Thus, we classified them according to the adopter categories proposed by Rogers (2003) using the available data. The distribution of our sample was: Innovators (T2); Early adopters (T4); Early majority (T1, T5, and T6); Late majority (T3), and Laggard (None).

The lesson plans proposed by those teachers were codified according to a previous classification developed to understand the maturity of technology use in the lessons. This classification was adapted from the Future lab (2014) framework. We divided this classification into four stages, as shown in Figure 3, labeled as "MAT" for maturity level. We did not consider the fifth stage of the framework due to the scope of this research, which focuses on the teacher. Thus, teachers were distributed as follows: Level 1 (T3); Level 2 (T1, T4, and T5); Level 3 (None), and Level 4 (T2, T6).
One determinant characteristic to classify these didactic sequences as level 2 is the use of technology to enrich existing approaches to teaching. For example, T4's sequence presented many characteristics of a level 3 plan, including higher-order thinking as his goal and involving students in a more collaborative and independent activity using the application. However, he used AR as a strategy to enrich more traditional approaches, which is a characteristic of level 2.

Based on the data collected, we have not found an evident relationship between the teacher profile and the level of maturity of technology use in the lesson plans. Nevertheless, some considerations can be made about this topic.

First, as expected, the innovator teacher created a lesson plan considered level 4. This teacher proposed more unexpected uses for the AR tool, including interdisciplinary content among her goals and a lesson centered more on students' independent development. T6 also planned a lesson considered level 4 in terms of maturity. Although considered early majority, she had previous experience using AR in her lessons. In these two lesson plans, it is noticeable an expansion of the tool's potential, which can be seen from the suggestions of adding texts and interaction between 3D objects and its use to expand students' skills and abilities.

Three lessons were considered level 2, two of those were proposed by teachers classified as early majority adopters and the other by a teacher considered an early adopter. T4, an early adopter, usually introduces a new technology with a familiar approach.

My intention is a practice I have as a teacher. When I'm using something new, I try to focus and not keep inventing too much because it gets a little out of control, right? If you're working on a new tool and want to get a very large set of features from it, or if you're working on a new tool and want to use other tools in the lesson, too, I think there's a lot for you to pay attention to and then you can lose a little bit of control of the pedagogical direction of the lesson…

As pointed out previously, the most determinant characteristic for this level is that teachers classified as level 2 tended to enrich their traditional teaching approaches. In our sample, the technology prototyped was new to most of these teachers. Only two teachers participated in previous phases of the research. The teacher classified as late majority proposed a lesson classified as level 1. He did not explore the full potential of the
technology. He planned to use it for students to create sentences using prepositions, the grammatical topic being studied, and illustrate them using AR.

Results have shown that the AR tool proposed has the potential to be used in different contexts and levels of maturity depending on the teacher's - and why not add the student's - creativity. The tool also allowed teachers to work with diverse skills and competencies, as evidenced in the lesson plans proposed. Results also suggest that teachers' unfamiliarity with the technology might generate a tendency to more conservative approaches and a need for control.

To summarize, as regards our first goal, data has shown that, in general, the VP Creator prototype allowed teachers to create their AR experiences proposed, although, in some cases, modifications were needed. The most visible issues were related to the limited 3D library and lack of 3D modeling possibilities.

Although teachers demonstrated interest in creating 3D content themselves, their speech revealed that this is not always easy due to a range of factors already discussed in this work, such as workload and knowledge. When asked about the possibility of content creation, T4 summarizes his ideas about this issue as follows:

Creating it himself. I think it would be interesting, although it would already be something that you would work to implement, knowing beforehand that it will be little used. That there will be a minority that will use it, right? But if it's possible to have it, I think it would be cool.

In his experience, some schools hire an employee to support teachers' use of technology and the creation of simple digital artifacts, such as videos and games. Also, the uses proposed seemed to vary according to the teacher's previous use of AR, and this aspect, along with teacher profiles and experiences with technology, also influenced the maturity of the uses proposed.

Finally, the code comments "COM" shown in Figure 3 captured the teacher's considerations about the concept or use of the tool. Most of them evaluated the concept positively, describing the tool as "cute and beautiful" (T1), "intuitive" (T2), "interesting" and a "very good and promising idea" (T3), and "cool" (T6). T2, though, mentioned that it could also have a high distracting potential. Thus, she suggested creating conceptual maps for students' stories in her lesson to guide them throughout the process.
Considerations About Virtual Playground Creator Interface (INT)

We observed if there were features missing or unnecessary in the VP Creator to adjust for the next iteration. When ranked by frequency, we conclude that most issues referred to difficulties understanding the interface (71 entries), as shown in Figure 3. Most of these comments were related to nomenclature such as "institution" and "private playground". For instance, T6 needed to register again because she got confused about the need to add the institution twice.

This type of result was expected since some of the functions were experimental and not completely implemented, such as the details for the lists and filters to find playgrounds. Another issue was related to how the teacher selects and organizes student groups. T2 had difficulties understanding where to click to see the comments. Adjustments regarding the interface were made to fix the issues raised, as detailed later.

The second most common code was suggestions for improvements and features (39 entries). The code justification of the suggestions (10 entries) was directly related to this code. Due to the exploratory nature of the test, teachers were encouraged to think about different possibilities with the application and, in some cases, expand and propose different uses for it that would suit their needs.

T2 proposed to add individual instruction feedback to students through the comment feature to avoid participation inequality in the groups. T5 demonstrated the desire to work with improvisation rather than a script. However, he recognized that with this idea, students would practice little reading and no writing. He emphasized that this option brings humor to the class through the unexpected sequence of the story. These particular suggestions were not implemented. The first one is because one of the goals of the tool is to promote group work, and thus, it would be easier for the teacher to have a centralized communication process. As for the second one, we understand it is up to the teacher's instruction rather than the tool itself.

Another frequent suggestion referred to the need to facilitate how teachers see what students will see as they interact with the application, which was accomplished through the addition of the button "view as student".

Other interesting suggestions by T6 were the addition of rubrics as well as the gamification of the experience so students would collect stars, for example, as they...
progress towards their goals. T1 suggested having the application for iOS. These suggestions have been collected for future iterations of the prototype.

Other suggestions involved the desire to highlight important points of the instruction to students as "students have the habit of scanning texts," according to T2. Both T2 and T5 suggested the use of avatars. As T2 put it: "When the student creates the character, it has that gamified motivational aspect of avatar creation that gives the impression of belonging to the student. So it's very, very cool."

Another important aspect was the integration of the AR tool to different steps of the lesson, both real and virtual. T2 suggested that students could take pictures of the mind map they create for the story and integrate it into the application. In her words:

I think it gives depth to what will be created in the tool. It ends up going beyond just creating a moment for the class, right? You can put the entire class, the whole line of learning reasoning, in the same tool, so it adds value to the tool and learning.

T5 also suggested using laminated images he has in his school as markers. These suggestions indicate the desire for integration between the AR application and the classroom environment, which is aligned with the literature findings (CUENDET, 2013).

The code lack of functionality or instruction had 26 entries. As regards functionality, T5 and T1 wanted to see their stories from a different angle, which was unstable due to marker issues. One way to solve that is to only provide more stable markers. Also, analyze the markers added by the users and let the tool tell them if the marker is stable and provide a replacement option if needed. Other issues related to instructions were the lack of clarity regarding what to do with the PIN number and markers, the lack of choice for the scenery marker, and the lack of a menu option in the student's module, among others. Most of these issues were addressed in the final updated version.

The compliments (25 entries) were related to different aspects of the tool, such as the possibility of interacting in the same room (T2) and the fact that the 3D objects are responsive (T3). T1 and T6 highlighted that they liked the comment option, which reinforces the importance of giving feedback to students. T5 pointed out that he liked the option to set the playground as private, whereas T4 highlighted the fact that the default
option is to leave the playground private. These comments reinforce the importance of moderating students' content.

Another compliment was related to the option topics to avoid, which was defined by T2 as "a positive limitation for the student". T3 mentioned that the adaptations made in the prototype comparing the first (test 1) and second versions (test 2) were cool. Most of these spontaneous compliments were related to pedagogical elements, such as promoting students' co-creation through instructions, monitoring their production (setting the playground as private or public), and providing feedback (through comments).

In general, the application was positively evaluated by the teachers. In fact, T5 stated that the application itself is a spectacle explaining that his suggestion for improvement, i.e. adding instruction throughout the story, could be done later and that it would be possible to use the application as proposed.

The code bug had 19 entries. Three teachers had problems registering for the application and inserting the groups, which made them close the application and start it again. That happened when they left one specific and optional field empty. The tool wrongly treated this field as mandatory, and since it was empty, it considered all the information invalid and did not save anything. After found, this bug was fixed.

T3 was not able to access the playground sometimes in the student's module. This issue occurred because he closed the application without logging out, which made the application not save the PIN. The teacher's recording of the use was important to verify this issue later. After that report, the authoring tool keeps the PIN, along with all other important data, saved after every change the user makes.

The code functionality does not achieve the goal has 7 entries. It refers to functionalities that do not achieve the intended goal or the expected result. The main issue was the question mark button, designed to offer help to the teachers, which in most cases did not catch their attention. They usually noticed it only after the researcher called their attention. T4 mentioned the following about it: "It's because I think I'm kind of a digital native. Because digital natives wouldn't want to know about question marks, right? Ah, I've seen it. I know what it is. I'll put right here what I think it is." Another topic was the instruction for the topics suggested to students. At first, T1 did not think she needed to be concerned with it, in other words, she did not understand its goal.
Finally, the code problems had 6 entries. This code referred to difficulties of use. For instance, T2 used her cell phone in dark mode and, thus, could not see some of the lines to fill out information while she was registering her group in the application.

Another problem that happened with T1 and T2 was that the keyboard covered certain buttons, leaving them confused about what to do next. After all, since we were doing the test online, teachers used their own devices, which allowed for a test in a less controlled scenario revealing problems that we would otherwise not be exposed to.

These results highlight the importance of the pedagogic aspects and instructions for teachers, helping them understand how the student will use the application to achieve the proposed learning objectives, as well as giving them some control over what will be produced and practiced. As T2 summarized: "We don't want creative control. We want the student to create freely, but security control has to be as much as possible."

**VP Playground Creator Update**

The UI changes in the second prototype was based on the participant’s feedback and suggestions as well as on the principles proposed by Peters (2014). The software used during the development of the prototype was Adobe XD.

The UI of the final application was designed as a digital functional prototype. As the prototype was being developed, there were some reviews during the process, aiming to avoid the majority of UI and UX issues, such as navigation problems, inappropriate icons and button sizes, and wrong animations.

Those reviews were made by design and educational professionals that used the application and pointed out the majority of details; focusing on a more intuitive experience, the navigation had, as a parameter, some default buttons, like return, home, settings, and profile. Figure 4 shows a few examples of the finished UI.

The navigability flow was designed focusing on fluid navigation that allows the user to always go back to the main menu, to the previous page, and also, access both the profile and settings screens. Some examples of the applications related to the design principles presented by Peters (2014), such as the feedback would be the response of the application to any kind of interaction of the user with it; the focus on cognitive goals is
seen in the whole experience of using the application; all playgrounds would have, on its
details, content descriptions and some teachers' extra notes; and others.

Figure 4 - Virtual Playground Creator final interface: (a) login screenshot; (b) my
playgrounds screenshot; and (c) instructions page.

During the development of the UI, some important points were reviewed and
corrections were made. Some of them are due to new reads in the user's test reports. Some
of those corrections were about the interface vocabulary, as some terms were not very
clear and created confusion among users.

Due to the vocabulary review, some interface sectors were removed, changed, or
merged with others. The private playground, for example; was a way for the teacher to
give some advice, comments, and corrections to the student's story. However, there was
a better way to structure that; using an interactive status screen as illustrated in Figure 5
(a), a chat screen as displayed in Figure 5 (b), in which both students and teacher can talk
in real-time and debate about any theme they need to.

The VP application shown in Figure 6 was designed to focus on an integrative
experience. The user does not need to exit the AR scene to look for any information about
the story. Also, we focused on designing a more minimalist and simple interface, using
intuitive icons and friendly navigation.
Students can create the story scenario as desired in their module. They need to choose an object or character as shown in Figure 7 and link it to a target; after, they need to take the printed target, place it on the chosen location, and use it in the story.

Figure 7 - VP Creator choice page on the student's module.
Additionally, the student has total control and a top view of the whole composition of the story. Just like the objects and characters mentioned above and the chapters and the scenes inside it as illustrated in Figure 8 (a) and (b), respectively.

Figure 8 - Story's timeline in the student's module of VP Creator.

![Figure 8](image_url)

Font: Image elaborated by the author.

Teachers have a whole view of the student’s story, as they can access the scenes and chapters. They can also access information about the group dynamics and statistics about the activity that allows them to monitor group participation (Figure 9). On this page, they might know the time and effort each student is investing in the project.

Figure 9 - Story's timeline in the student's module of VP Creator.

![Figure 9](image_url)

Font: Image elaborated by the author.

**CONCLUSION**

Literature review has shown numerous advantages of the use of AR technology, such as increased motivation and cognitive performance. Moreover, there are many applications currently available for educational use. Nevertheless, its use is still not
widespread in education as most applications were created with little or no partnership with teachers or students.

Thus, this work presented the improvements of an AR application and its authorship tool, VP and VP Creator, focused on language learning, developed collaboratively with teachers, designers, developers, and researchers. We contributed to the literature by presenting: the improved prototypes of both tools conducted through cycles of DBR research with a multidisciplinary team of an educational researcher, and computer science as well as the evaluation of VP with teachers, which validated the application and helped to improve it as well as to determine further steps. The second round of tests presented in this paper was the base for their refinement.

It is noteworthy that the context of the COVID-19 pandemic imposed changes in all contexts of life and consequently in education, including the second round of tests carried out during this research. As expected, the participant teachers experienced very different realities in the context of the second test, involving fully remote or online teaching, hybrid teaching as well as face-to-face teaching, which provided a broader richness to our data.

As regards our first research question, i.e, whether VP Creator allows teachers to create AR experiences to use in their classrooms, results have shown that the tool allowed teachers to work with diverse skills and competencies as evidenced in the didactic sequences proposed. In general, the VP Creator prototype allowed the teacher to create their AR experiences proposed, although, in some cases, modifications were needed. The main issues were related to the limited 3D library in the prototype evaluated and the lack of possibilities for creation, which answered our second research question, that is, whether there were missing or unnecessary features in VP Creator. The input received was used to improve the UI and UX of both VP and VP Creator.

Although no clear relationship could be established between the teacher's profile and the maturity of the didactic sequences proposed, results suggested that the teacher's unfamiliarity with the technology might generate a tendency to more conservative approaches and a need for control. The maturity of the lessons proposed was still not very high.

Results have also reinforced the importance of the use of statistics to aid teachers in the assessment process and improve student learning, which could also encompass
different types of data including multimodal learning analytics (BLIKSTEIN, 2013); as well as the importance of flexible assessment including the use of rubrics and peer assessment. Additionally, results have suggested the need for integration between AR experiences and other classroom activities.

As we refine the prototype of both VP and VP Creator, we intend to test them in a real-life scenario with students to fully understand their potential as well as to improve these tools.

Finally, we understand that our small sample size might have skewed the results. Authors are aware that the teachers who used AR are in a way pioneers in the use of technology and may not represent the general population. Thus, the findings cannot be generalizable.

ACKNOWLEDGMENTS

Authors would like to thank the participants and Fundação de Amparo à Ciência e Tecnologia de Pernambuco (FACEPE) (processes IBPG-0605-1.03/15) for their funding as well as the participants.

REFERENCES


POCRIFKA, D. H.; CARVALHO, A. B. G. *Ciclos de codificação e o software atlas.ti: Uma parceria criativa para análise de dados qualitativos em pesquisas sobre o uso...


Submetido em 10/01/2023.

Aprovado em 22/03/2023.