Journal on Interactive Systems, 2026, 17:1
ISSN: 2763-7719 • doi: 10.5753/jis.2026.6605

RESEARCH PAPER

Lessons Learned in Human Evaluation: A Case Study of Augmented Reality in Books

Bruno Adachi [Federal University of Ouro Preto | *bruno.adachi@aluno.ufop.edu.br*]
Júlia Gonzaga [Federal University of Ouro Preto | *julia.gonzaga@aluno.ufop.edu.br*]
Paula Fernandes [Federal University of Ouro Preto | *paula.cintra@aluno.ufop.edu.br*]
Sílvia Bim [Federal University of Technology – Paraná | *sabim@utfpr.edu.br*]
Silvio Boss [Federal University of Technology – Paraná | *silvioboss@utfpr.edu.br*]
Saul Delabrida [Federal University of Ouro Preto | *saul.delabrida@ufop.edu.br*]
Natasha Valentim [Federal University of Paraná | *natasha@inf.ufpr.br*]

Departamento de Computação, Universidade Federal de Ouro Preto, Rua Professor Paulo Magalhães Gomes, 122, Bauxita, Ouro Preto, MG, 35400-000, Brazil.

Abstract. Books adapted for Augmented Reality (AR) combine traditional printed pages with digital layers of interactive content such as 3D models, animations, or audio, accessed through mobile devices, creating a hybrid reading experience that merges physical and virtual elements. This paper presents a case study on the use of AR in printed books, focusing on the challenges and lessons learned from human-centered evaluation. Using the children's book "Alan Turing: His Machines and His Secrets" as a case, we analyzed the feasibility of integrating AR features into works not originally designed for this purpose. The study involved testing two AR Software Development Kits (SDKs), Vuforia and ARCore. Results indicated that Vuforia offered the most efficient image recognition, supporting 71% of page detections when the smartphone was handheld and 76% with device support. A usability assessment with 18 participants, based on the System Usability Scale (SUS), revealed positive acceptance of AR interactions but also highlighted ergonomic limitations when handling books and mobile devices simultaneously. The findings emphasize the importance of experimental design, pilot testing, and consideration of user habits in AR evaluation studies. Overall, this work provides practical insights to guide future illustrators, developers, and researchers in adapting books to AR contexts.

Keywords: Augmented Reality, AR, Books, Human Evaluation, Usability

Edited by: Elisa Rodrigues | **Received:** 24 September 2025 • **Accepted:** 30 June 2026 • **Published:** 11 July 2026

1 Introduction

Technology is increasingly present in daily life, facilitating simple tasks such as communication to more complex activities such as bank transactions. In this way, technological advances expanded, and several spheres of society had to develop to keep pace with advances, especially education.

In Brazil, the Base Nacional Comum Curricular – BNCC (Common Core Curriculum) [BRASIL, 2018] was approved in 2017 for Kindergarten, Elementary, and Middle School and in 2018 for High School. The BNCC is a reference document that determines the competencies (general and specific), skills, and essential learning that all Brazilian students must develop during each stage of primary education. The document is unique and seeks to reduce educational inequalities in Brazil.

Among the competencies to be developed, we highlight the fifth competence: Digital Culture, which aims to "Understand, use and create digital information and communication technologies in a critical, meaningful, reflective and ethical way in the various social practices (including school one) to communicate, access and disseminate information, produce knowledge, solve problems and exercise protagonism and authorship in personal and collective life"[1] [BRASIL, 2018].

In this way, many technologies have been inserted into the routine of students and teachers in order to facilitate learning and promote student interest. Whether through a new teaching model such as hybrid teaching or tools that allow more excellent monitoring of the development of each student, as well as educational applications and a virtual learning environment.

However, the insertion of technology alone only partially contemplates the competence of Digital Culture. Thus, seeking to contemplate the demands arising from this competence, in October 2022, the document Computação - Complemento à BNCC (Computing - Complement to the BNCC) [BRASIL, 2022] was approved. All these documents result from several years of elaboration and discussion with different sectors of society, including, in the case of Computing, the Brazilian Computing Society [Ribeiro *et al.*, 2023].

Faced with the approval of these documents, one of the challenges is teacher training and the production of didactic material. In this way, a series of children's books was launched in August 2022 to present historical computing personalities and their contributions to the field. The first title in the series presents the biography of the father of Computing, Alan Turing [Boss and Bim, 2022].

Using Augmented Reality (AR) mobile applications, we can achieve a low-cost solution to expand students' learning experience, making classes more interactive and enjoyable. "These technologies have the power to extend learning experience to new places which are distant from the classroom" [Abazi-Bexheti *et al.*, 2022].

Thus, this paper aims to present a study on the use of AR in printed books, exploring both technical and human

[1]The English translation of this excerpt from the BNCC is the authors' own.

aspects of its adoption. For instance, we implemented 3D animations and interactive objects that appear when specific illustrations of the book are scanned with a smartphone. We investigated the feasibility of applying AR to a children's book not originally designed for this purpose, evaluating the performance of different AR Software Development Kits (SDKs) and conducting usability tests with real users. By combining technical analysis with human-centered evaluation, this work identifies the main challenges, opportunities, and lessons learned when integrating AR into educational materials. The results contribute to a better understanding of how AR can enhance the learning experience, guide future developments in educational resources, and support the broader goals of digital culture in the Brazilian educational context.

It is important to clarify that this current work is an extended work encompassing two previous papers. Them being: Augmented Reality in Books: An Evaluation of Alan Turing Book [Adachi *et al.*, 2024], and Lessons Learned in Human Evaluation: A Case Study of Augmented Reality in Books [Adachi *et al.*, 2025].

This manuscript is a substantially extended version of our previous works, which separately addressed the technical feasibility of AR SDKs for book integration and methodological lessons from conducting human-centered evaluation. In contrast, the present study integrates technical performance analysis with complete empirical usability results, providing statistical comparisons, cross-analysis of six different ergonomic configurations, and consolidated design guidelines. This unified and expanded approach enables broader and more generalized insights that were not possible in the earlier fragmented publications, thus constituting a substantive and original contribution.

This paper is organized as it follows: Section 2 details why it is important to approach this problem, Section 3 presents some key concepts to better understand the work, Section 4 presents some works related to the topics addressed in this paper, Section 5 describes how we conducted this study, Section 6 presents the results from the SDKs tests, Section 7 presents the results from the usability evaluation, Section 8 details the discussions and lessons learned through the evaluation, and Section 9 summarizes the conclusions.

2 Problem Statement

In the following subsections, we will discuss the importance of the use of technology in schools and how it can contribute to a better learning experience. As well as the importance of evaluation with real users.

2.1 Relevance of Technology for Teaching

Nowadays, technology plays a fundamental role in people's daily lives, and this is no different in the case of education. Using different technologies in classrooms offers a series of benefits that can transform the way content is taught and how it is received by students [Demitriadou *et al.*, 2019].

Through computers and mobile devices, teachers and students can access information quickly, ensuring a wide range of educational materials at any time, enriching the learning experience. In addition, tools such as educational games can make the experience more dynamic and engaging, encouraging more active participation by students [Freitas, 2018].

The aspects raised previously also favor a more personalized learning experience, where content is taught at the pace and based on the needs of each student, making classrooms a more friendly and inclusive environment.

Bittencourt and Albino [2017] argue that education in the 21st century must respond to the profile of digital natives and to the structural transformations lead by Information and Communication Technologies (ICT). The authors emphasize that the pedagogical use of digital media should promote interaction, autonomy, and critical thinking, rather than merely replacing traditional tools.

Similarly, the study on technology as a learning tool proposed by De Barros [2019] highlights that technological resources can enhance student engagement and support the teaching-learning process when integrated intentionally into pedagogical practice. Both works reinforce that the relevance of technology in classrooms derives not simply in access to devices, but in its strategic and reflective incorporation to promote meaningful learning experiences aligned with contemporary social and cognitive realities.

However, as digital ecosystems evolve, it becomes necessary to deepen investigations into emerging technologies, such as AR. This technology introduces immersive, interactive, and spatial learning capable of enhance students' visualization, contextualization, and experiential understanding of complex content. Despite its growing presence, there remain a need for empirical and theoretical studies that examine its impact in learning environments.

Despite the widely documented benefits of technology integration in classrooms, the literature also highlights important limitations and risks associated with its use. A study conducted with teacher, administrators, and parents in elementary schools in Kosovo indicates that, although technology encourages students to acquire new knowledge and supports contemporary teaching practices, it may also generate significant drawbacks [Kryeziu *et al.*, 2021]. Among the most frequently reported concerns are increased distraction, reduced social interaction, feelings of loneliness, unequal access to technological resources, and privacy risks.

These findings suggest that the educational impact of digital technologies is not inherently positive, rather it depends on pedagogical intentionality, proper monitoring, infrastructure conditions, and continuous teacher mediation.

2.2 Relevance of Human-Centered Development

User-centered development uses multiple investigative methods such as interviews and analysis to understand user needs. There are several approaches in the literature that can be used according to each context. In general, the approaches aim to fulfill the following steps: 1) Understand the context of use; 2) Specify the requirements; 3) Design solutions; and 4) Evaluate [Interaction Design Foundation, 2016]. This process can generally occur with the direct participation of end users in all or most of the steps through Participatory Design [Cumbo and Selwyn, 2021] or Collaborative Design [Friesen and Jacobsen, 2021] approaches. Or, users may be involved only in the evaluation phase.

It is important to highlight ethical aspects of research involving human beings. Although each country may have legislation with considerations based on their contexts, these must be guaranteed according to local legislation concerning the people volunteering in the study.

The user-centered development process is relevant because all development time is focused on the user, which opens up opportunities for identifying requirements and solutions, improving usability and user experience once the product is ready.

3 Theoretical Background

The subsections bellow highlights some key concepts to better understand this work and the possibilities provided by the technologies we used.

3.1 Extended Reality (XR) and Its Technologies

Extended Reality (XR) [Mann *et al.*, 2018] is a term that represents technologies that allow you to extend a person's concept of reality by combining virtual elements with the real world or an immersive world[2]. To experience XR, we use technologies such as Virtual Reality (VR), Augmented Reality (AR), and other emerging technologies. According to Skarbez *et al*. [2021] we can describe these technologies as it follows:

- VR: creates an entirely virtual and immersive environment without overlap or interaction with the real world. It can be accessed through Head-Mounted Displays like VIVE and Oculus. Some even include a "head tracking" feature, a sensor that allows the user to execute rotational and translational movements in virtual reality, composing six degrees of movement.
- AR: following the Reality-Virtuality Continuum, AR is positioned closer to the real environment, as it superimposes virtual elements onto the physical world while preserving direct perception of reality.

Three principal alternatives exist to interact with real-world context and AR virtual entities according to Chen *et al*. [2019]:

- Marker-Based (Images): Uses image recognition to identify objects already programmed into an application or device.
- Contour-Based (3D): uses visual codes to detect and position virtual objects in a physical space.
- Location-Based (GPS): positions virtual objects in space using GPS data.

3.2 Augmented Reality Software Development Kits

A Software Development Kit (SDK) can be described as an all-in-one solution, providing developers with tools containing all of the essential components for development, like compilers and libraries [GitHub, 2025].

The main benefits of using a SDK are that it offers support to multiple platforms, making it easy to develop and deploy an application on multiple devices and operational systems. Another benefit worth mentioning is that they provide some ready-to-use pieces of software, allowing developers to incorporate them into their projects and reducing time and money costs.

The following subsections presents the leading AR SDKs.

3.3 Vuforia SDK

The Vuforia SDK is an AR platform for application development. This SDK enables the creation of AR applications with support for most phones, tablets, and extended reality glasses.

It enables developers to create applications that combine virtual and real-world elements for an immersive experience and uses computer vision technology to recognize and track planar images and 3D objects in real time. So, Vuforia SDK is widely used to create AR experiences that realistically interact with objects and real-world environments.

The platform provides a tool that permits the developers to check how well it will track targeted images, rating them on a scale of 1 to 5 stars. According to its documentation [Vuforia, 2026] a 4 or 5 rating brings the best results regarding detecting and tracking images.

The documentation also provides a best practices guide reporting on what makes a good image target. There are four attributes that we need to take into consideration. The first one is about the image features. A feature can be a sharp, spiked detail. The image analyzer identifies these details as small yellow crosses. The accuracy is better as more crosses are identified. Figure 1 shows an example of features identification.

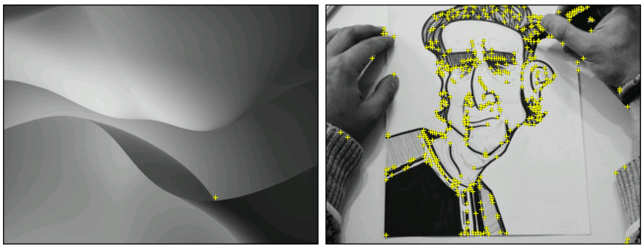


Figure 1. Features comparison, the one on the left is an image with low number of features compared to the right image. Retrieved and adapted from [Unsplash, 2026].

The second attribute refers to the image's contrast. That means the elements in the targeted image must contain an apparent difference between them. Figure 2 presents a practical example of good and bad contrast.

Figure 2. Contrast comparison, the one on the left is an image with a low level of contrast compared to the right image. Retrieved and adapted from [Unsplash, 2026].

The third one refers to the distribution of the features

[2] According to Dede [2009], the sense of immersion on virtual reality is directly related to the user's willing suspension of disbelief, which is the result of sensory, actional, and symbolic elements. So, the immersion in a digital environment is a singular experience for each person.

on an image. If an image contains, for example, more yellow crosses on the left portion and just a few ones on the right side, the image rating will decrease. Figure 3 shows an image with poor distribution of features.

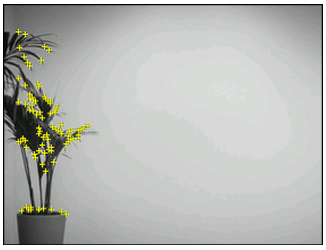


Figure 3. Example of an image with poor features distribution. Retrieved and adapted from [Unsplash, 2026].

The last attribute that can interfere with the rating level is repetitive patterns. Even if the image contains many features and a good contrast level, repetitive patterns can hinder the detection performance. For example, an image of a chessboard will get a bad score because a chessboard is composed only of black and white squares. Figure 4 presents an image with repetitive patterns.

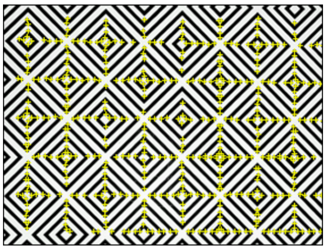


Figure 4. Example of an image with repetitive patterns. Retrieved and adapted from [Unsplash, 2026].

3.3.1 Vuforia Features

The Vuforia SDK can track many artifacts and categorize them as images, objects, and environments. Among its main features, the SDK provides:

- For images and objects: Vuforia allows tracking images or markers to create a real AR experience.
- For environments: Vuforia allows the detection of flat surfaces to make the AR experience more realistic.

3.3.2 Vuforia Tools

An SDK provides ready-to-use tools that can incorporate another application. The list below describes two of them.

- Vuforia Model Targets: A tool for allowing developers to create AR applications that recognize and track 3D objects.
- Vuforia Ground Plane: A tool for allowing developers to create AR applications that can detect ground plane surfaces and accurately adjust the position of virtual objects.

3.3.3 Vuforia Engine API

The Vuforia Engine is an Application Programming Interface (API) integrated into Vuforia's SDK. The API offers a wide range of easy-to-use features, such as tracking images and customizing cameras, and also allows integration of Vuforia's Engine with other libraries.

3.3.4 Vuforia Engine in Unity

Vuforia Engine in Unity integrates Vuforia's AR Engine with the Unity Game Engine. The integration enables developers to create AR applications using Unity as their development platform. They make game development much easier, as professionals can use a single platform to create AR and game elements.

3.4 ARCore SDK

ARCore SDK is an AR platform developed by Google. It is designed to help developers to create AR applications for Android and iOS devices. ARCore uses computer vision, sensors, and machine learning technologies to create an accurate and stable experience. Its three main features are:

- Motion Tracking: ARCore uses data from sensors such as an accelerometer, gyroscope, and compass to track the device's movement and provide a stable experience.
- Environmental Understanding: ARCore uses cameras and some computer vision techniques to detect the size and location of flat surfaces (e.g., floors, tables, and carpets) in the environment and use them as a basis to create AR objects.
- Light Estimation: By estimating the lighting conditions of the actual environment, ARCore uses the device's camera and sensors to adjust the illumination of AR objects so that once generated, the illumination of the object is similar to the natural environment.

Like Vuforia's SDK, ARCore provides the user a tool to check the target image performance and a best practice guide documentation [Google, 2026]. According to its documentation, a tool called "arcoreimg" can determine an image score according to its detecting and tracking potential. Unlike Vuforia, which rates the target images on a star scale, ARCore gives the image a score between 0 and 100, with 75 being the minimum score recommendation.

Developers should consider some features when selecting reference images. The image resolution must have at least 300 pixels of width and length, and it is essential to note that increasing image resolution does not improve the overall performance. Alongside three features must be avoided: the usage of highly compressed images because they can interfere with the feature extraction process, the usage of images that contain elements with very few features (e.g., bar codes, QR codes, logos, and other types of line art), and images with repeating patterns. These factors can interfere with the final score of the image target.

3.5 ARKit SDK

ARKit is Apple's unique AR platform for iOS and iPadOS devices. It is a set of software tools that allows developers to create AR applications for iPhones and iPads. These tools are listed below:

- ARKit Tracking: allows the device to track the user's movement to provide a stable AR experience.
- Plane Detection: used to detect flat surfaces in the environment and use them as a basis to create AR objects.
- Light Estimation: used to adjust the illumination of AR objects to match the lighting conditions of the natural environment.

- Object Detection and Image Tracking: used to detect and track specific images to display AR objects.
- Face Tracking: used to track faces and create and apply special effects.
- Scene Understanding: used to understand the environment and enable the creation of interactive AR objects that react to the environment.
- Integration with Other Apple Technologies: used to integrate with other Apple pieces of technology, like the virtual assistant Siri, enabling developers to create more advanced AR applications.

Through ARKit's official documentation [Apple, 2026] developers can learn more profoundly all the tools mentioned before. As our interest is in image tracking, we searched through the documentation to find out how to achieve good performance.

Firstly, we learned that, when providing the reference images, the user must provide the image's physical size as accurately as possible because ARKit uses this information to determine the distance between the image and the camera.

Some other factors that can interfere with the detection process are the use of nonplanar surfaces (e.g., bottles, labels) and the lack of consideration about different lighting conditions, as glossy papers and device screens can cause reflection.

It is also worth mentioning that Apple's tool for developing applications, Xcode, can provide a quality estimation of the reference image and present helpful information, like the lack of contrast between elements.

3.6 Comparative SDK Features for AR Book Context

This section discusses the main differences among the SDKs and establishes two of them that present as the most promising for the objective of our work. As seen before, the three SDKs work similarly, so we have considered that there is no point in testing all of them. So, we have selected some criteria to deepen our discussion about them.

The chosen characteristics are: 1) image target support; 2) which types of devices are supported; 3) pricing; 4) what development platforms are available; 5) if they require an internet connection all of the time; and 6) how they measure the target images. Table 1 summarizes our discussion.

We first checked if they could detect and track plane images because our tests involved using a mobile device camera to capture images from a book. One of our primary concerns is how we can reach a broader audience with our application. Then, we took into consideration two criteria relative to accessibility. First, we researched how many devices and systems each SDK could support, and after that, we checked if the platform could work without an always-on connection requirement.

The Vuforia and ARCore SDKs were the platforms that are able to cover the majority of devices; ARKit, on the other hand, only supports two types of devices: the iPhone and iPad. The obligation to have an internet connection all of the time to use an AR application could affect the ability to reach more people. Due to social issues, many people in Brazil do not always have an internet connection.

The last two elimination criteria were about the easiness of development and maintaining infrastructure. The last one did not present an issue because all of them were free or presented with a free-of-charge plan. While Vuforia and ARCore SDKs can reach a broader range of development platforms, Apple's SDK is available only for Apple's devices.

After our evaluation, we concluded that ARKit presents some problems regarding accessibility and easiness of development. As our goal is to reach a broader audience, we decided to continue our work only with Vuforia and ARCore SDKs.

4 Related Works

The following subsections explore the current usage of AR applications combined with physical books and works related to the usability aspects of AR technologies.

4.1 Augmented Reality for Books

The work conducted by Dias *et al*. [2019] proposed a case study to evaluate AR SDKs in classrooms. The authors gathered a set of SDKs with AR technology based on markers. They developed a mobile application to raise which characteristics of a classroom impact performance and which SDKs proved to be more efficient based on their objectives. The work developed in this article used techniques similar to those proposed by the authors but for evaluating book pages. This work gave us directions on how we could evaluate different AR technologies as part of our methodology for this project.

A study by Cheng [2017] explored books with AR capability among 153 university students (composed of 94 undergraduate students and 59 graduate students). The author aimed to investigate the students' perception of cognitive load, motivation, and opinions about using such books through questionnaires.

The stimulation of reading since childhood can bring numerous cognitive benefits. However, considering the early age, difficulties related to the recognition of new words can impact the comprehension of the story and discourage the practice of reading. Based on these problems, Danaei *et al*. [2020] proposed a study to measure and compare text reading comprehension using a traditional and printed book with AR capability.

Researchers have empirically identified shared reading between parents and children as beneficial to children's language and literacy development. In addition, emerging technologies such as AR can bring even more benefits, but what is the parent's perception of the technology in question? The work by Cheng [2019] aims to explore issues related to User Experience regarding the use of AR. The work aims to summarize the parents' perceptions, expectations, and intentions about using traditional books combined with AR technology.

Lai *et al*. [2015] explored the usage of AR combined with physical books to make the students' learning experience more interactive. They developed a smartphone application using Unity and Vuforia to track and detect images of an educational book. Their goal is to evaluate the students' opinions about the use of AR on smartphone surfaces. Thirty-seven invited students tested the application and post-responded to a survey with two simple questions. The first one was, "Do you agree that the application can help you to understand the

Table 1. SDKs comparison: Vuforia, ARCore, and ARKit.

SDK	Supports Image Target	Supported Devices	Pricing	Development Plataforms	Internet Required	Rating Scale
Vuforia	Yes	iOS, Android, UWP, Eyewear Devices	Free without the usage of cloud services	Android Studio, Unity, Xcode, Visual Studio, Lumin	No	1 to 5 stars
ARCore	Yes	iOS, Android, Android Device Emulator	Free	Android Studio, Unity, Unreal Engine,WebXR	No	0 to 100
ARKit	Yes	iOS, iPadOS	Free	Xcode	No	Not Found

book's content?". The answers were quite positive, and 34 students strongly agreed with the affirmation. The second question was, "Do you agree that the application can make the learning more interesting?". Also, they mainly received positive responses, with 35 students strongly agreeing with the affirmation.

Dong and Si [2018] conducted a similar study where the author created an AR application to be used alongside a children's book to make the learning experience more enjoyable, interactive, immersive, and realistic. The chosen book was about dinosaurs, and the first step that the authors took was to create multimedia based on the content of the book. These multimedia included audio, video, and 3D models.

Another worth mentioning work is the MyVision AIR: An Augmented Interactive Reality Book Mobile Application by Al-Ali *et al*. [2016]. The main goal of the work was to promote book reading among adults by enhancing the experience through an AR application. Using Unity and Vuforia, the authors could include content in a book based on the images contained in it. These contents included 2D images to provide additional information, 3D models of significant national and international landmarks, animations, virtual buttons to trigger certain events, and translation of certain parts of the book.

4.2 Augmented Reality Applications

Zambiasi and Pozzebon [2025] examined the acceptance of Interactive AR technology virtual fitting rooms in the Brazilian fashion e-commerce context through an extended Technology Acceptance Model (TAM). By analyzing data about Perceived Enjoyment (PE), Perceived Informativeness (PI), and System Quality (SQ), the study shows that hedonic and informational factors play a more central role than purely utilitarian beliefs in shaping user responses. The findings show that PE is the strongest predictor of Attitude (AT) and also significantly influences Perceived Usefulness (PU), whereas PU does not directly affect attitude, diverging from the traditional TAM assumptions. Overall, the study concludes that, in immersive retail technologies, experiential and informational dimensions outweigh strictly functional attributes reinforcing the need to adapt TAM to interactive AR context.

Mazon *et al*. [2022] report the implementation of the AR Facilitating Science Teaching (RAFEC) course, designed to train public school science and biology teachers in the use of AR as a complementary educational tool. The study highlights AR's potential to facilitate the understanding of complex and abstract concepts through interactive 3D visualization, even in contexts with limited laboratory infrastructure. The authors conclude that AR enhances the teaching-learning process and reinforces the need for continued teacher training to support its effective integration into basic education.

4.3 Gestures in Augmented Reality Applications

AR applications are not limited to visual overlay, they are fundamentally interactive systems in which learning and task execution depend on how users manipulate digital elements anchored in the physical environment. In mobile AR applications, interaction typically occurs through gesture-based input, including touch gestures (e.g., tap, drag, pinch, rotate), device movement, and spatial repositioning. These gestures operate as the primary interface layer between the user and the augmented content, shaping how information is explored, controlled, and interpreted. Consequently, the design and evaluation of gesture-based interaction play a decisive role in determining usability, engagement, and overall user experience in mobile AR contexts.

Dong *et al*. [2022] conducted a study to compare two different types of interactions used in AR applications. The first type of interaction was through touches on the smartphone screen, while the second involved gestural movements in the air using the smartphone and its sensors. The study included various smartphones interactions and applications and also conducted user testing.

It is worth noting that although the study focused on AR, the researchers did not use AR applications exclusively because of the limited availability of such software in application stores. As a result, it was necessary to develop a custom application to conduct a more comprehensive study. It is important to note that the authors of this study did not include the usage of of AR applications combined with physical books. The main goal of citing this paper is to give an overview of how developers and users think and interact with these types of apps.

The related work demonstrates the relevance of using AR in books, mainly through the findings based on the user studies. Any found-related work evaluates the integration of the current technology and a book written not for AR use, and this is an identified opportunity covered by this work.

In the study by Dong *et al*. [2022], a range of applications were analyzed based on recommendations from the Google Play and Apple App stores. Finally, six applications were selected to present a good representation of gestures commonly used in games, painting, design, and education apps. Twenty-one participants between 18 and 59 years old were recruited for the study. Table 2 presents the applications chosen for the research and the gestures that participants should execute.

After listing the applications, the participants watched videos that explained how to interact with the apps. Researchers requested them to imagine two different forms of interaction. The first involved using only their fingers and the mobile device screen, while the second required movements in the air with the device. Finally, the users rated the

Table 2. Selection of apps and gestures.

#	App	Gesture(s)
1	Angry Birds	Slingshot
2	Pokémon Go	Throw
3	Just a Line	Draw Erase
4	ARia	Open Drawer Close Drawer
5	IKEA Place	Move Rotate
6	CooolAR	Scale Down Scale Up Open Door Close Door

gestures on three criteria: Goodness-of-Fit, Ease-of-Use, and Engagement.

At the end of the experiment, the researchers collected 504 gestures and notes related to them. To analyze the results, the authors reviewed the participants' notes and evaluated the agreement level between the suggested gestures. Based on this analysis, a set of gestures was defined, as shown in Figure 5.

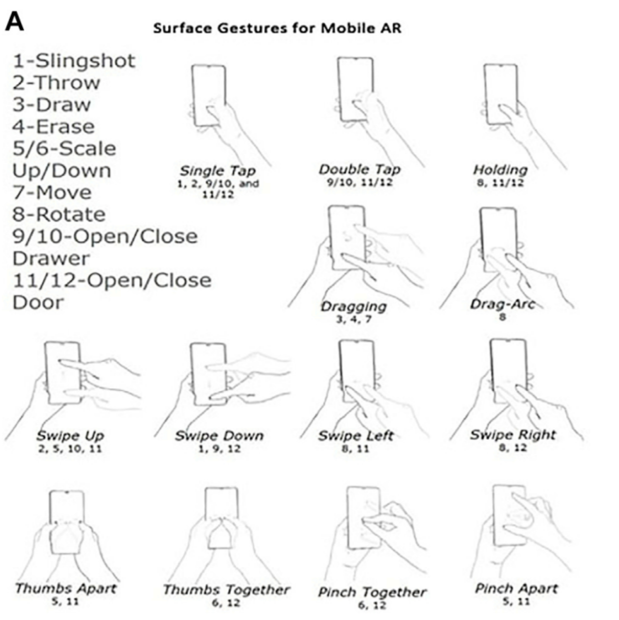


Figure 5. Set of surface gestures-based interaction with mobile AR applications. Retrieved and adapted from Dong *et al*. [2022].

5 Methodology

The Figure 6 below shows illustrates the overall design of this work.

Firstly, we conducted a test regarding the viability of using the AR SDKs discussed in Section 3 to observe how well they can handle the tasks of detecting and tracking illustrations of an already published physical book that was not designed to be used with AR technologies. It is represented in the Figure 6 as the first block, and it will be detailed in Subsection 5.2.

During the test we noticed some usability and ergonomics problems. The act of holding the book with one hand and aiming the smartphone camera with the other was found to be cumbersome. As these problems could affect the perception and adoption of this technology, we proposed to conduct an in-depth investigation on how we can overcome these obstacles. To execute this experiment we developed an AR smartphone application (discussed in Section 5.4) and tested using the methodology better explained in Subsection 5.3.

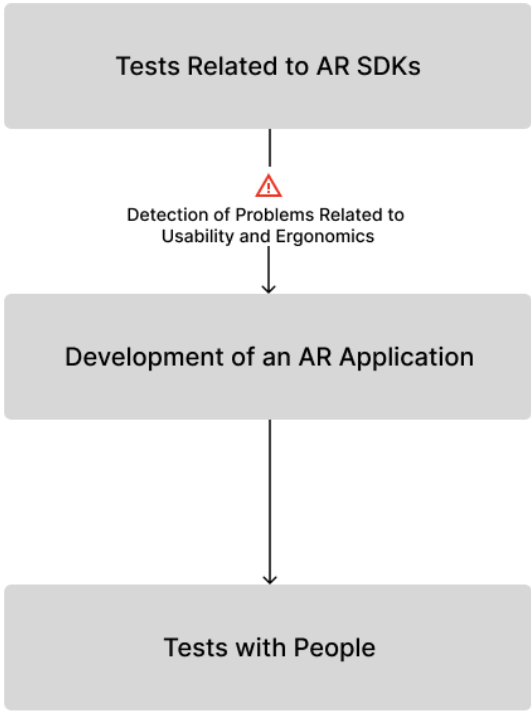


Figure 6. Methodology diagram.

5.1 Material

Considering the demand for teaching materials that can somehow support the teaching of Computing in Basic Education, the children's book "Alan Turing: His Machines and Secrets" [Boss and Bim, 2022] was selected for the research described in this work. Although commercially the illustrated book is aimed at children aged 8 to 12 years old, using AR resources can potentially expand this audience.

In addition, we could expand the use of the technology, covering disciplines such as Portuguese, History, Geography, Physical Education, and Foreign Languages in addition to Computing itself. The use of AR can also offer accessibility features, allowing people with visual or hearing impairments, for example, to have access to the book's content.

The book's authors are two Computing professors and researchers with works on Computing Education, Human-Computer Interaction, and Computing Theory. The book presents the biography of Alan Turing and brings a didactic example of the Turing Machine for the sum of 1 + 2. The series has the seal of the Brazilian Computing Society and the Digital Girls Program [Maciel *et al*., 2018].

5.2 Tests of AR SDKs

The chosen book [Boss and Bim, 2022] has 64 pages in total. First, we extracted 63 pages in image format to evaluate them using the image analyzers of the chosen SDKs. One of the pages was excluded from our tests because it was blank. To carry out our tests we've used an Samsung Galaxy M51 mobile device, which has the following specifications: a Snapdragon 730 chipset, 6GB of RAM and four rear cameras (64 Mp + 12 Mp + 5 Mp + 5 Mp).

As mentioned in Section 3, the chosen SDKs provide tools capable of evaluating an image according to its detection and tracking capacity. For this, they attribute a grade to these images, and the scale used by Vuforia ranges from 0 to 5 stars, and ARCore uses a numerical scale ranging from 0 to 100. According to the documentation of the SDKs, the recommendation is to use images rated at 4 and 5 out of 5 stars for Vuforia and a minimum score of 75 out of 100 for ARCore.

The first test, carried out using image analyzers, aims to verify the detection and tracking capacity of the book in its current state. For evaluative purposes, we will decompose the results into three categories: images that obtained the result recommended by the documentation, images that obtained a medium rating, and images that obtained a low rating. According to the Vuforia SDK documentation, a low rating refer to images that obtain a 1 or 2 stars score, so we have considered a 3-star score to be a medium rating.

Images that fit in the low rating category will be subjected to image manipulation techniques to improve their performance. In the medium rating category, we will carry out performance tests using a mobile application to capture the elapsed time in order to find out how bad in terms of detection and tracking capabilities the use of these images can be.

5.3 Usability Evaluation

The experiment carried with people had as its main reference a previous study done by Dong *et al.* [2022], where the authors evaluated user interaction with AR applications to verify people's perception of different forms of user interaction with such applications. Although the mobile application market offers a wide range of AR software, most of them still use a form of interaction focused on 2D applications, through touches on a surface (in this case, the mobile device's screen). The authors then question whether interaction through movement gestures, using the devices' sensors, might be more intuitive and more relevant to AR applications.

Although the AR application was developed based on a children's book, the usability evaluation was intentionally conducted with adult participants. The objective of this study was not to assess children's learning outcomes or educational effectiveness, but rather to investigate the usability, ergonomics, and interaction challenges associated with combining a physical book and a mobile AR application. Adult participants allowed us to identify interface, interaction, and handling issues while avoiding the additional ethical, methodological, and recruitment complexities involved in studies with minors. Since the evaluated aspects concern the interaction aspects rather than age-specific learning processes, adults were considered an appropriate population for this stage of the research.

To conduct the experiment, we developed an application using the Unity tool combined with the Vuforia SDK, which worked as follows: when pointing the smartphone camera at an illustration from the book we chose as the object of study, a 3D object appeared (as can be seen in Figure 7) and, with that, the evaluation volunteer had to execute an interaction based on a series of tasks (the same as shown in Table 2). The creation of this application had the objective to facilitate the evaluation, since, in the study by Dong *et al.* [2022], different applications were selected to test different tasks and gestures.

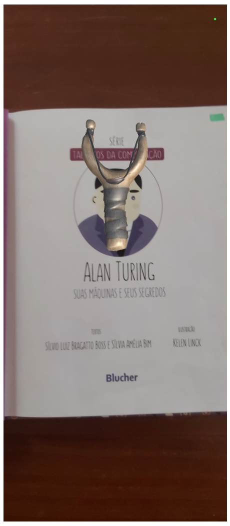


Figure 7. Example of page detection and augmented object.

Each evaluator had to execute 12 tasks in six different scenarios, totaling 72 tasks by each evaluator. In the following paragraphs we are going to describe each scenario used in the evaluation.

Hands Free: the evaluator must manipulate the book with one hand to turn the pages, and the other must be used to interact with the smartphone (as shown in the Figure 8a).

Fixed Support: the smartphone is fixed in a support (as shown in the Figure 8b) while the evaluator can manipulate the book with both hands. The support allows the user to change its angle.

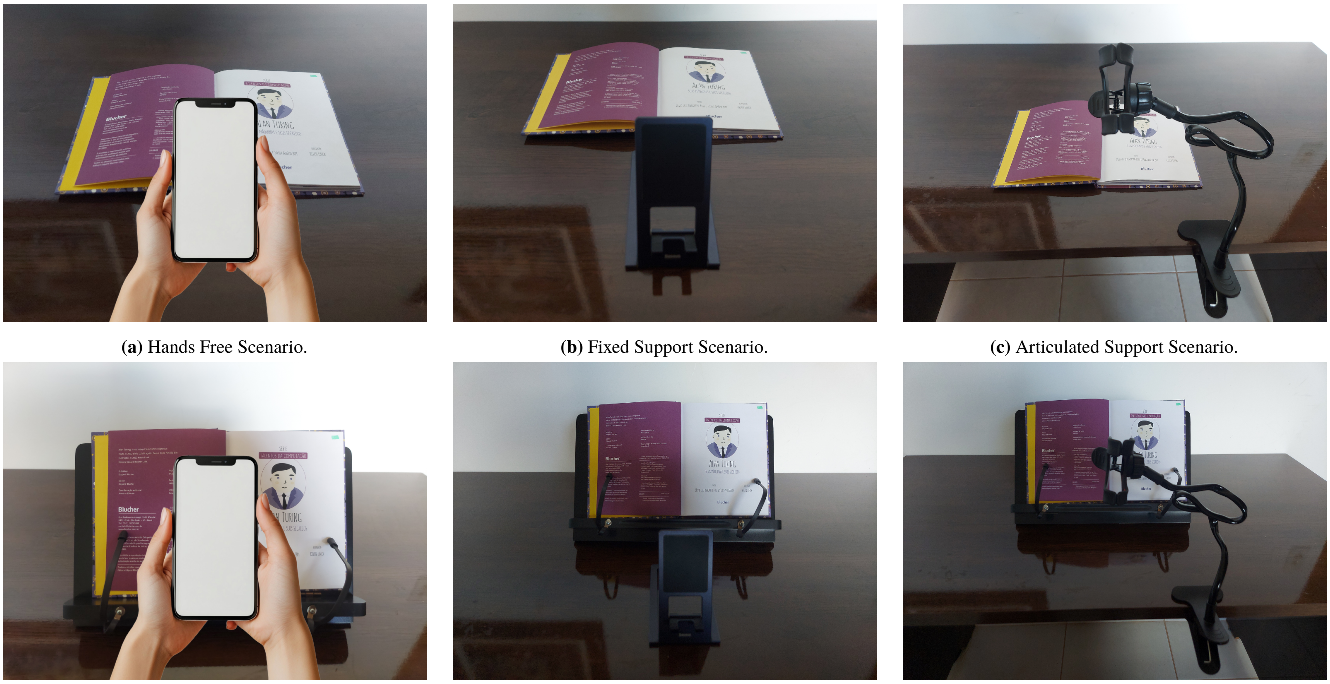
Articulated Support: the smartphone is held by an articulated support (as shown in the Figure 8c), allowing the evaluator to change the smartphone height and angle, while both hands are free to manipulate the book.

Hands Free and Book Support: the book is held by a support which the evaluator can switch its height and angle, while both hands are free to interact with the smartphone (as shown in the Figure 8d).

Fixed Support and Book Support: the smartphone is fixed in a support while the book is held by a support (as shown in Figure 8e). Allowing the user to user both hands to manipulate and adjust the supports or to interact with the smartphone.

Articulated Support and Book Support: the smartphone is held in an articulated support while the book is held by a support (as shown in Figure 8f). With both hands free, the evaluators can adjust the supports individually and interact with the smartphone freely.

After each scenario, the evaluators were asked to fill a System Usability Scale (SUS) [Brooke, 1995] form. And at



(a) Hands Free Scenario. **(b)** Fixed Support Scenario. **(c)** Articulated Support Scenario.

(d) Hands Free and Book Support Scenario. **(e)** Fixed Support and Book Support Scenario. **(f)** Articulated Support and Book Support Scenario.

Figure 8. Scenarios evaluated in the usability test.

the end of the evaluation we asked the evaluators to fill a Technology Acceptance Model (TAM) [Ma and Liu, 2005] form. During the whole evaluation we encouraged the evaluators to give us a feedback regarding each scenario.

The affirmation contained in the SUS Forms are listed bellow:

1. I think I would like to use this system frequently.
2. I find the system unnecessarily complex.
3. I found the system easy to use.
4. I think I would need the help of a person with technical knowledge to use the system.
5. I think the various functions of the system are well integrated.
6. I find the system to be very inconsistent.
7. I imagine that people will learn to use this system quickly.
8. I found the system cumbersome to use.
9. I felt confident using the system.
10. I had to learn a lot of things before I could use the system.

5.4 A Tool for User Gestures Interaction Evaluation of Augmented Reality in Books

Up to this point, the work has focused solely on testing the performance of specific SDK's in recognizing pages from a particular book. We have not yet incorporated any form of user interaction into the application we developed, an essential aspect of our study that we have yet to address. This is mainly because our initial project did not include such an investigation within its scope.

Throughout the development of this work, it became apparent that ergonomic considerations for handling the physical book and mobile application are crucial for technology adoption. Therefore, a comprehensive study of how users interact with AR applications is necessary to ensure successful adoption.

We developed the application using the Unity game development tool and the Vuforia SDK, which, as we already discussed, produced the best results in detecting book pages in previous tests. The application has three sections: a menu screen, three introduction screens, and a section for detecting pages and interacting with 3D objects.

In the study conducted by Dong *et al*. [2022], different applications were used to test the set of gestures. We decided to develop a single application for all the gestures for ease of testing.

To test the slingshot gesture, we have used a matching asset where the objective is that the user drags the elastic of the slingshot to the appropriate gesture. As for the throw gesture, a basketball and a hoop were used, capturing the essence of the Dong *et al*. [2022] chosen application Pokémon Go.

A whiteboard was used to draw and erase the gesture. A cubic shape was used for the move, rotate, scale down, and scale up gestures. Moreover, we have used a drawer and door asset for the open drawer, closed drawer, open door, and closed door.

Figure 9 shows the application menu. The screen displays the laboratory's logo along with two buttons: one to start the application and another to display more information about it.

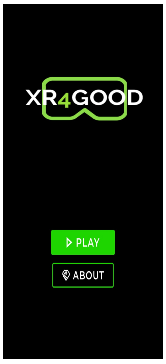


Figure 9. Menu screen of the proposed mobile AR application.

Figure 10 displays the three screens for introducing the application[3]. They intend to onboard users by explaining the application's purpose and what needs to be done.

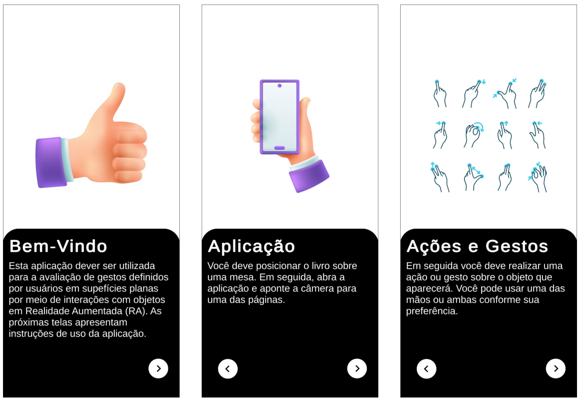


Figure 10. Onboarding screens.

Figure 11 presents three examples of detection from three pages of the book Alan Turing: His Machines and Secrets and the display of 3D objects.

Figure 11. Detection of different pages of the book an their augmented objects.

In addition to the three previous examples, the system contains five other objects. Thus, we contemplate the 12 gestures proposed by the study mentioned above. Table 3 presents all the objects created and their respective forms of interaction.

[3]The text presented are in Portuguese-BR, and their translation are: "This application must be used to evaluate user defined gestures on flat surfaces by interacting with AR objects. The next screens show instructions on how to use the application", "You must place the book on a table. Next, open the application e point the camera at one of its pages", "After that, you must execute an action or gesture over the object. You can use one or both hands."

Table 3. Developed 3D objects and their interactions.

Page	3D Object	Interaction	Page	3D Object	Interaction
1	Slingshot	Slingshot	6	Cube	Scale Up
2	Basketball and Hoop	Throw	7	Cube	Move
3	Whiteboard	Draw	8	Cube	Rotate
4	Whiteboard	Erase	9	Drawer	Open/Close
5	Cube	Scale Down	10	Door	Open/Close

5.5 User Study Design

In this study, a combination of quantitative and qualitative methods was employed to evaluate the AR application from multiple perspectives. The SUS was used to obtain a standardized, quantitative measure of overall usability. SUS is widely recognized for its simplicity and reliability, providing a single composite score that reflects users' perceptions of ease of use, effectiveness, and satisfaction with the system. Its brief format allows for quick administration while still yielding meaningful insights into the usability experience.

Complementing this, the TAM was applied to assess users' perceptions of Perceived Usefulness (PU). TAM provides a theoretical framework to understand how these factors influence the intention to use the system, offering a structured approach to evaluate the potential adoption of the AR application. By measuring PU, researchers can identify which aspects of the system contribute most to its acceptance and highlight areas for improvement.

To enrich these quantitative assessments, the Think-Aloud Protocol was employed, in which participants verbalize their thoughts while interacting with the application. This method captures qualitative insights into users' cognitive processes, revealing usability issues, misunderstandings, or strategies that are not apparent through questionnaires alone. The Think-Aloud data thus complement the SUS and TAM results, providing a detailed understanding of the user experience and informing iterative improvements to the system.

By integrating SUS, TAM, and Think-Aloud, the study combines objective measures, theoretical assessment, and direct observation to deliver a comprehensive evaluation of the AR application, addressing both usability and user acceptance.

5.6 Ethical Conduct

Conducting evaluations involving the participation of volunteers must be very well structured, and all risks must be considered [Del-Masso *et al.*, 2014]. Therefore, before starting the experiment, a series of documents were written and evaluated by the Ethics Committee of the institution in question (in this case, the Universidade Federal de Ouro Preto), following the legislation on research involving human beings in Brazil.

Among the documents evaluated by the Ethics Committee, there are two that are extremely important: the Research Brochure and the Evaluation Project. The Research Brochure is a document that contains information about scientific research and is used by the committee's evaluators to verify the validity of the study that will be conducted. The Evaluation Project is a normative document that must describe the main information about the experiment to be conducted:

- The project design (how the evaluation will take place);
- The hypothesis that the authors intend to answer;
- The primary objective;

- The methodology (the evaluation methods to be used);
- Data related to the recruitment of volunteers (criteria for inclusion and exclusion);
- The risks to which participants will be subject;
- The benefits of participating in the research;
- And sample size.

The research documents reported in this article were then submitted to the Ethics committee of the Universidade Federal de Ouro Preto. The research was approved under the number of Certificado de Apresentação de Apreciação Ética (CAAE[4]) 76992024.2.0000.5150.

In accordance with the approved ethical protocol, all collected data were anonymized to remove personally identifiable information. The anonymized dataset is publicly available in a scientific repository to promote transparency, reproducibility, and educational reuse. The dataset is available in Section 9.

The entire evaluation took place in a testing laboratory within the premises of the Universidade Federal de Ouro Preto (as can be seen in Figure 12). A specific area of the room was used and another area related to evaluations with Virtual Reality technologies was isolated.

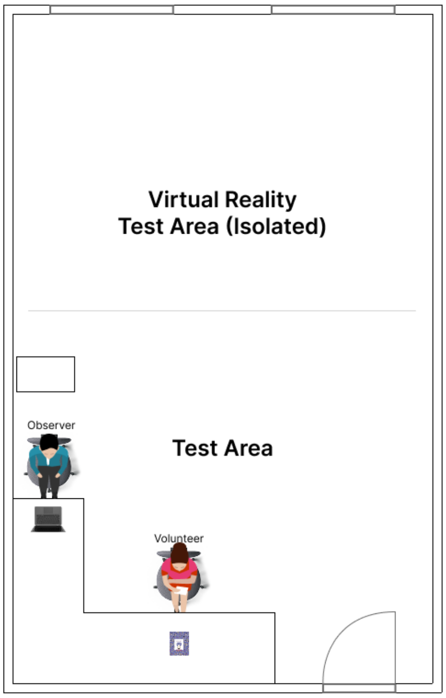


Figure 12. Top view of evaluation room.

The area where the experiments were carried out is part of a room specifically designed for user evaluations, a private space where noise and distractions are avoided. In the case of this study, only the observer and the volunteer were allowed in the room, so that any embarrassment on the part of the volunteer would not be an obstacle, and they could express themselves freely.

The positioning of those involved was also designed with the aim of avoiding possible embarrassment on the part of the volunteer, since direct eye contact could, in some way, inhibit their behavior, and important information for the study could be disregarded.

It is important to emphasize that the observer's point of view allowed all actions performed by the volunteer to be observed, facilitating note-taking.

The next step was to call for volunteers to carry out the experiment; the main method used to do this was sending messages to online student communication groups. At the end of the evaluation, 18 volunteers contributed to the study.

The volunteers were welcomed and the evaluation protocol began, where the reason for conducting the study was explained and how their participation would help the research. Next, the Termo de Consentimento Livre e Esclarecido (TCLE[5]) was presented, which contained relevant information about the evaluation, such as the risks presented and the possibility of the participant withdrawing.

Considering the case study work and how the results would be examined in order to create a better understanding of the problem, it was agreed between the authors to collect data on the experiment volunteers, data observed during the experiment and a general assessment by the volunteers of the proposals presented to try to solve the problems addressed.

Immediately after the acceptance of the volunteer to contribute to the study, the evaluation began. In order to carry out a thorough analysis of the results, we considered it necessary to collect some information about the participants, such as age, gender, level of education, course (if the person was enrolled in higher education), and whether the volunteer was already familiar with AR technologies.

After collecting the sociodemographic data, an order in which the tests would be conducted was chosen. This was necessary because the evaluation included a total of six different scenarios, and if we started all the evaluations with the same scenario, factors such as fatigue could affect the results obtained. Each scenario used different equipment, so when starting a new scenario, the objects were presented to the volunteers, as well as the possibilities of interacting with them. The following objects were used: a fixed and an articulated support for smartphones, and a book support. By combining these objects, we tested the following scenarios: hands-free (the volunteer held the smartphone with one hand and the other held the book), fixed support (the smartphone was fixed to the support), articulated support (the smartphone was held by the support), hands-free and book support, fixed support and book support, and finally, articulated support and book support.

During the evaluation, volunteers were asked to execute a series of tasks and, for each of them, a gesture had to be chosen. And, using the Think Aloud methodology, volunteers also made comments about the scenario in question; this data was filled in a spreadsheet using a laptop by the observer. At the end of each scenario, each person was asked to fill out a SUS form and, upon completing all six scenarios, a TAM form and was also asked for each task what their opinion was about the best scenario. The evaluation was then finished.

[4] Certificate of Presentation of Ethical Appreciation

[5] Free and Informed Consent Form

6 Results of AR SDKs Tests

This section presents the results obtained by the tools of the ARCore and Vuforia SDKs. An overview of our results can be seen in the Figure 13.

In addition, images that achieved 3-star performance in Vuforia were evaluated in terms of performance with a tool developed by the authors.

6.1 Results of Vuforia SDK

Using Vuforias' image analysis tool on the 63 images, we arrived at the following results: none of the images were rated 0 stars, seven images (11%) were rated 1 star, six images (9.52%) were rated 2 stars, twelve images (19%) were rated 3 stars, seven images (11%) were rated 4 stars, and thirty-one images (49%) were rated 5 stars. Figure 14 illustrates the results obtained.

Of the 63 images, 38 (60%) obtained grades equal to 4 or 5 stars. Such images have good detection and tracking performance, meaning that more than half of the book is suitable for using AR technology.

6.1.1 *Images with 1 and 2 Stars*

We selected all the images that obtained a score equal to 1 or 2 stars and cropped them to achieve better results. The use of this technique aims to verify if, when we focus on some specific element of the image to eliminate possible irrelevant information for the detection, the algorithm's performance increases. Figure 15 shows an example of one of the cropped images. In this case, we used the upper part. We discarded the lower part as it only contains texts with repetitive patterns, which may impair recognition and tracking as specified by the manufacturers.

Figure 16 presents the results: of the 13 images, four achieved an increment in the score, one image went from 2 to 5 stars, one image went from 2 to 3 stars, one image went from 1 to 4 stars and one image went from 1 to 2 stars. The rest of the images obtained an equal or worse result.

The mixed results indicate that using this image manipulation technique must be evaluated for each case. However, we have some indications of factors contributing to the increase or decrease in scores.

Figure 17, presents an image that had its score increased from 2 to 5 stars and an image that had its score decremented from 2 to 0 stars. Both images were from the Vuforia analysis tool. The '+' symbols represent the points identified by the algorithm as a traceable pattern according to the suggested attributes. The image is in black and white because color is not a meaningful property for Vuforia's algorithm.

(a) Score Increment **(b)** Score Decrement

Figure 17. Increase and decrease on image score. Retrieved from Boss and Bim [2022].

In the image on the left, there is a more significant number of attributes, a clear contrast between the elements, and a better distribution of attributes in the image. We attribute the increase in score to these factors. As for the image on the right, there is a smaller amount of attributes, a barely visible contrast between the elements (the book cover at the top blends into the background of the image), and there is a poor distribution of attributes, where most of them are at the bottom of the image. We attribute the decrease in the image score to these factors.

Since we want to maximize the traceability potential of the book, there are other options besides keeping images with low scores. Therefore, we evaluate the rating of the same image when we add a QR code to it. Figure 18 illustrates the edition that promoted increased attributes and a better distribution of them. Because the QR code is black and white, it was possible to enhance the image's contrast. In this way, the image evolved from a score of 0 to 4 stars.

6.1.2 *Images with 3 Stars*

In the image analyzer test, we selected all the images that obtained a 3-star score to observe the actual impact on performance that such a score could bring. Our objective was to understand the impact that these images can have in an actual activity of using the book.

To carry out this test, developing an application for mobile devices using the Vuforia SDK was necessary. In this application, we point the camera at the images with a 3-star score and measure the time the SDK takes to detect the image. Figure 19 shows our application pointing to one of the book's pages.

The sphere in the center of the page represents an augmented virtual object after recognition. We kept the sphere as the default object after recognition. However, we assigned a different color to each of the 13 images to check if the SDK could have confused the pages. For example, given the similarity of a page P and or P', the SDK can recognize P' when receiving P as input.

We conducted 30 tests for each of the 13 images. Among them, 15 used both hands to point the smartphone camera at the image, and the other 15 used a fixed support for smartphones. We realized that in most cases, the support reduced detection time. The support facilitates the image's framing process, eliminating problems with camera focus and poor stabilization, which usually happens using hands.

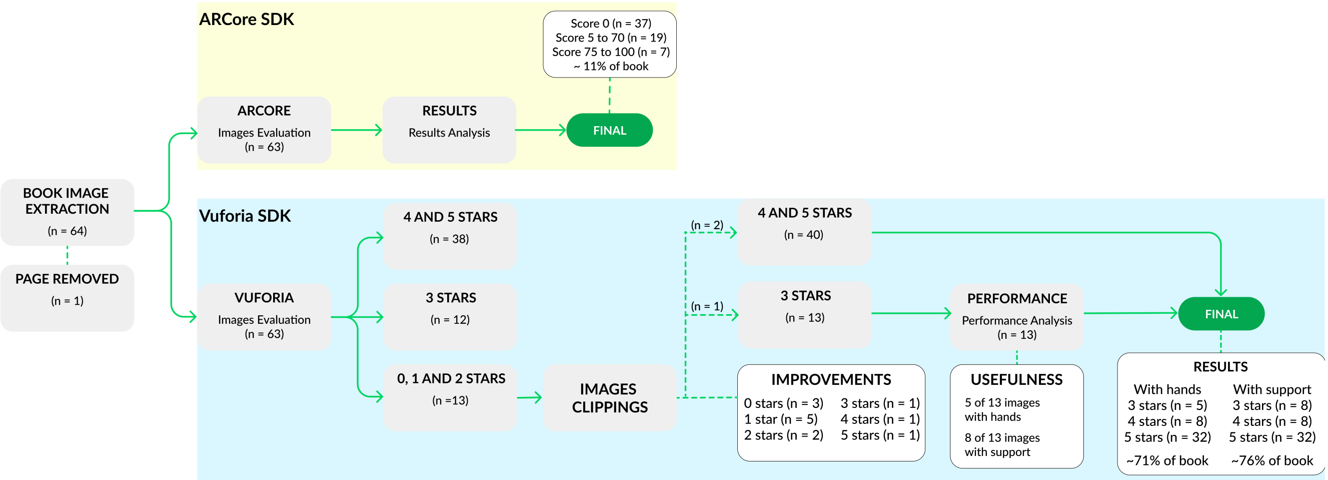


Figure 13. Alan Turing book evaluation procedure and results.

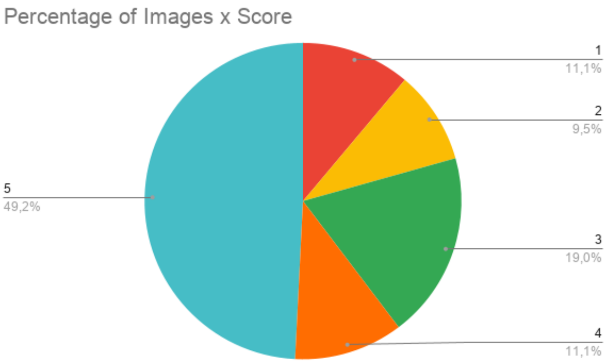


Figure 14. Scores on Vuforia SDK.

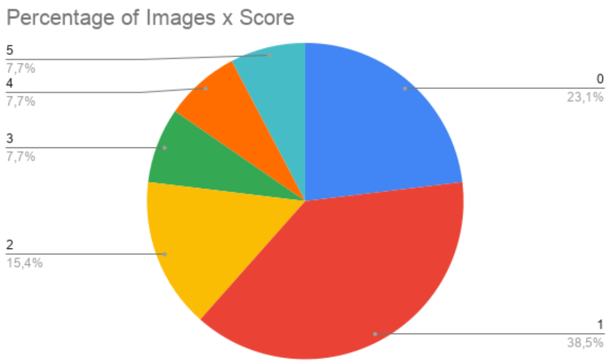


Figure 16. Scores of cropped images on Vuforia SDK.

Figure 15. Cropped image. Retrieved from Boss and Bim [2022].

Figure 18. Cropped image with the addition of a QR code. Retrieved from Boss and Bim [2022].

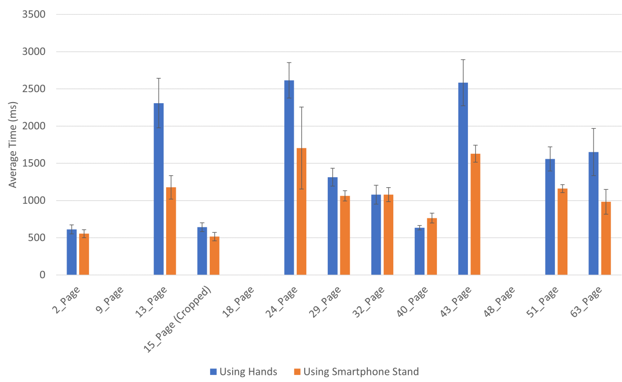


Figure 20. Average detection time and standard error.

Figure 20 presents the average time (in ms) of detection for each image, the standard error with a confidence interval of 95% for each image.

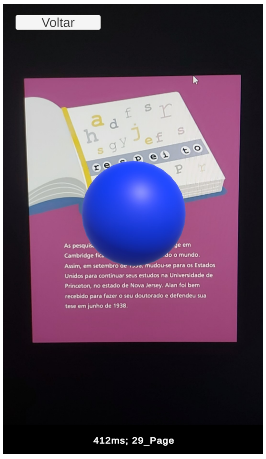


Figure 19. Application developed using Vuforia SDK to detect different pages an the elapsed time.

We empirically determined a maximum time of five seconds acceptable for image detection. If an image was not detected, it was considered a recognition failure. Three of the 13 images tested were not recognized in any of the 30 tests performed. Therefore, they do not show values in the Figure 20 graph.

Pages 24, 43, and 63 showed at least one recognition failure when the hands held the mobile. The same was repeated only on page 24 when the test happens with the mobile stable on support. For this reason, the average times of these images are longer and have a more significant variance, consequently increasing the error. Note that the average detection time of the other pages tends to be less than 1.5 seconds when using the fixed support. Only page 13 showed statistical significance with recognition time above 1.5 seconds when the hands held the mobile. The authors' perception is that this time of 1.5 seconds is acceptable for waiting in the process of recognizing a page.

6.1.3 General Results

In general, we classify pages into two groups. Acceptable and not acceptable for AR use. We consider that with the Vuforia SDK, 10 of the 63 pages (15% of the pages) analyzed cannot recognize the way they are in the book and would need a re-editing of the image or inclusion of some elements that can be used, such as a QR code. These 10 pages are in the unacceptable group. Of the remaining 53 pages, 40 (63% of pages) are well recognized with 4 or 5 stars, rated by the Vuforia SDK, and require no intervention. Of the 13 remaining images with 3 stars, we included eight of them in the acceptable group considering that they are recognized in an average time of up to 1.5 seconds and in the case that the tabletop support stabilized the mobile camera. In this scenario, the utilization becomes approximately 76%. In the scenario where we have the tests holding the mobile with their hands, the amount of usable images drops to five additional pages, generating use of approximately 71% of the book's pages.

6.2 Results of ARCore SDK

The tests repeat with 63 images taken directly from the chosen book [Boss and Bim, 2022] for the ARCore SDK. However, the results of this SDK are far below expectations for the book used.

Using the ARCore image analysis tool, we found that: 37 images (58%) scored 0, one image (1.58%) scored 5, four images (6.34%) scored 10, five images (7.93%) scored 15, four images (6.34%) scored 20, one image (1.58%) scored 30, one image (1.58%) scored 40, one image (1.58%) scored 65, two images (3.17%) scored 70, one image (1.58%) scored 75, one image (1.58%) obtained a score of 80, and finally, five images (7.93%) obtained a score of 100. Figure 21 presents the graphs with the results obtained previously mentioned.

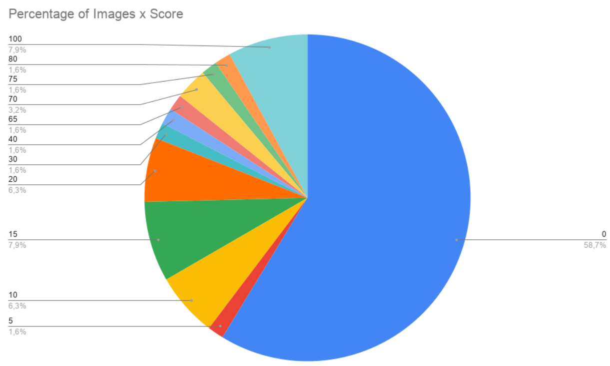


Figure 21. Scores on ARCore SDK.

For this project and case study, using the ARCore SDK proved to be unfeasible since, considering the total number of pages in the book, 58% obtained a score equal to zero. Due to this fact, we agreed it was not worth to use the same techniques used before (image cropping and adding a QR Code) because, in a real scenario, this would be expensive and time-consuming, escaping from the scope of our work.

Regarding the images that obtained scores equal or above 75, we could observe:

- There are no textual elements on these pages in question;
- Illustrations have a more significant number of attributes;
- The illustrations have more details than the others, and the attributes are better distributed on the pages. That is, there are few empty spaces.

The Figure 22 shows an example of image and its attributes that achieved a score above 75 on the ARCore image analyzer.

It is important to emphasize that ARCore proved unfeasible for the evaluated book [Boss and Bim, 2022], but this does not mean that it is an inferior tool or has a worse algorithm. For books with different styles of art and illustration, the SDK in question can get better results.

Due to the low quality of the initial results presented by ARCore, we chose not to carry out additional tests using the tool since the results show evidence that, for the case study, the tool needs to be adequate.

7 Results of the Usability Evaluation

The following subsections presents the social demographic data collected from the volunteers of the evaluation and the results from the SUS Forms for each scenario.

7.1 Social Demographic Data

The evaluation had the participation of 18 volunteers. As a form to help us better analyze the results from the SUS

Figure 22. An example of image that achieved a score above 75.

Form, at the beginning of each evaluation we collected some information about the volunteers, these being: age, gender, education level, course and if they had previous experiences regarding AR technologies.

The majority of volunteers had an age ranging from 18 to 29 years old (77.8%), 16.7% ranging from 30 to 39 years old and 5.6% ranging from 40 to 49 years old. As for the gender 72.2% of the volunteers identified as males and 27.8% as females.

Related to the education level of the volunteers, 83.3% are undergraduates, 11.1% has an incomplete postgraduate studies and 5.6% has a complete postgraduate studies. As observed, all of the volunteers are part of the academic community, but ranging across multiple courses.

Computer Science students represents 27.8% of all the volunteers, following by Pharmacy students (22.2%), Control and Automation Engineering is tied with Mining Engineering with both represented by 11.1% each. Physics, Law, Architecture and Urbanism, Geological Engineering and Industrial Electrical Engineering all represented 5.6% of the volunteers, each one.

As for the level of previous experiences regarding AR technologies, 61.1% of volunteers answered that they had prior experience with this type of technology, while the other 38.9% reported no prior experience.

7.2 Interpreting the Results

For each scenario tested, we created a graph (Figures 23, 24, 25, 26, 27, 28), and they should be interpreted as it follows: the X axis represents a score on the Likert Scale and the numbers on the Y axis represents an affirmation (as shown in Subsection 5.3). The numbers represents the level of consensus of the 18 evaluators (e.g., in the Figure 23, for the first affirmation: "I think I would like to use this system frequently.", zero evaluators gave a 1 score rating, two evaluators gave a 2 score rating, two evaluators gave a 3 score rating, seven evaluators gave a 4 score rating and seven evaluators gave a 5 score rating).

It is important to point that in a SUS form, the affirmatives corresponding to an odd number (1, 3, 5, 7 and 9) are formulated in a positive way, so scores closer to 5 represents a good level of satisfaction. On the other way, affirmatives corresponding to an even number (2, 4, 6, 8 and 10) are formulated in a negative way, in this case scores closer to 1 are preferred.

We then calculated the overall score, using the SUS Score Equation (ranging from 0 to 100):

$$SUSScore = ((A - 5) + (25 - B)) * 2.5$$

Where A represents the sum of the scores for all odd-numbered affirmations. And B represents the sum of the scores for all even-numbered affirmations.

7.3 Hands Free

The Figure 23 presents the results obtained for the Hands Free scenario.

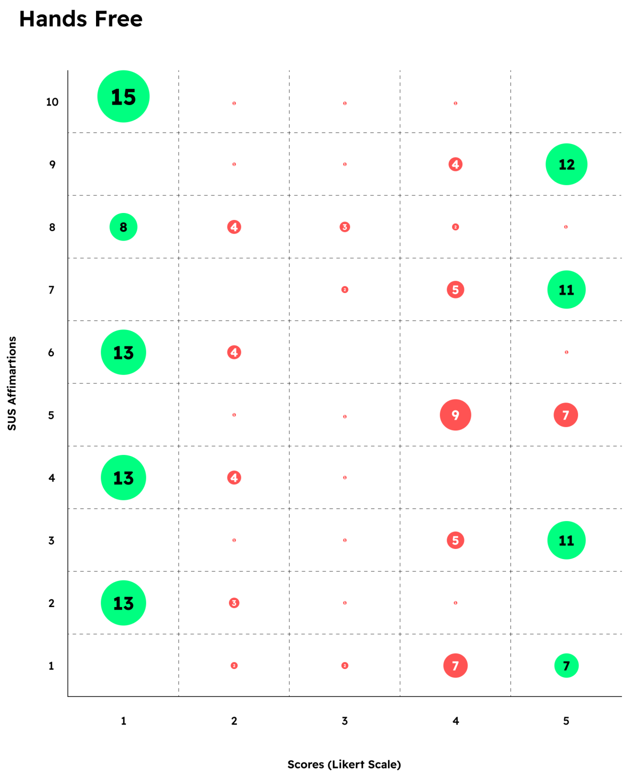


Figure 23. Hands Free scores.

We applied the equation for the scores given by each evaluator and then we calculated the average. The Hands Free scenario achieved a score of 85.14. According to Søegaard [2024] a score above 70 represents a good level of usability and over 85 represents an excellent level of usability.

As we can observe, there was an overall strong level of consensus among the evaluators. Although the average score was high, we noticed that two evaluators gave very low ratings regarding the Hands Free Scenario. With the scores given by the evaluators 5 and 7 the SUS Score obtained was 65 and 37.5, respectively, well bellow the suggested one.

Reading the notes we took during the evaluation we observed common complaints from both evaluators. Evaluator 5 (male, 19 years old, Computer Science student, with AR previous experience) said as it follows: "Performing actions with both hands is difficult". And evaluator 7 (male, 19 years old, Physics student, with AR previous experience) gave us the following insights: "It is hard to keep the book open" and "It is hard to interact with both the application and the book at the same time".

7.4 Fixed Support

The Figure 24 shows the results obtained in the Fixed Support Scenario.

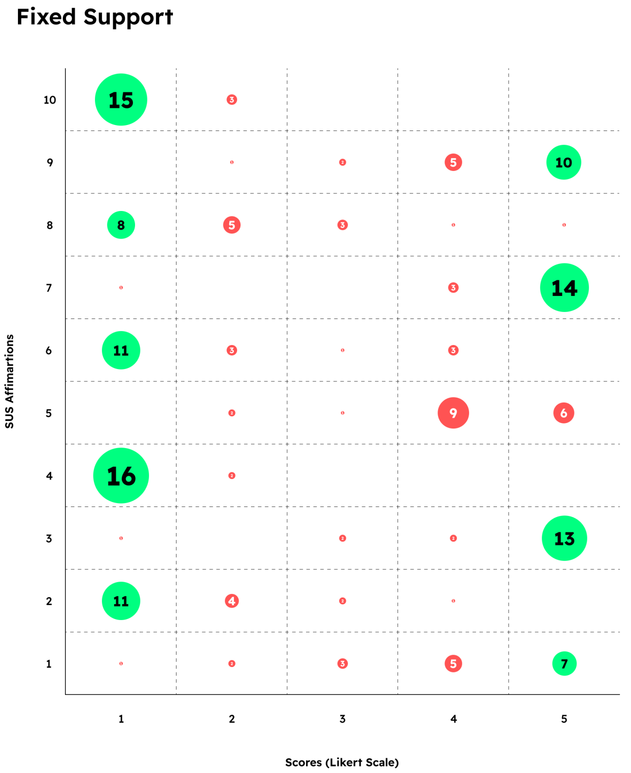


Figure 24. Fixed Support scores.

Using the same equation to calculate the average SUS Score, this scenario achieved a 84.03 score. As observed, there was also an overall strong level of consensus among the evaluators.

For this scenario, we observed that four evaluators gave ratings that, when calculating their individual SUS Scores, did not achieve the suggested score of above 70. The ratings given by evaluators 2, 5, 11 and 14 achieved the following scores: 60, 70, 32.5 and 70, respectively.

Based on the notes taken, we can conclude that the Fixed Support provided a better stability, which was helpful in the act of pointing the smartphone camera towards the book, and allowed the evaluators to manipulate the book with both hands facilitating the interaction.

But since it is a fixed object, there were reports of issues related to the height and angle provided by the support (which only the angle can be adjusted). Evaluator 14 made the following statement: "The Hands Free Scenario is more comfortable because holding the smartphone is less tiring than holding the book".

7.5 Articulated Support

The Figure 25 shows the results obtained in the Articulated Support Scenario.

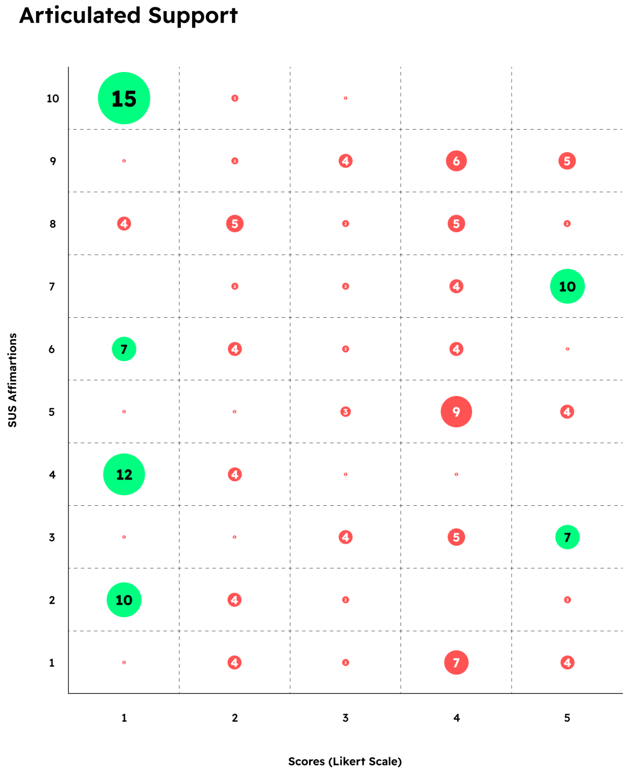


Figure 25. Articulated Support scores.

Despite achieving a 73.33 average SUS Score, indicating a good level of usability there was a decay in the evaluators level of consensus compared to the previous scenarios. Among the 18 evaluators seven of them did not achieved a score rating above 70 when calculating their individual SUS Score.

Evaluators 1, 2, 3, 9, 10, 14 and 18 achieved the following scores: 55, 42.5, 12.5, 67.5, 65, 60 and 52.5, respectively. Among the reports, issues related to the instability of the support were mentioned by 5 of the evaluators. Also, four of them claimed that the use of the support was beneficial because both hands were free to manipulate the book, but as soon as they needed to make small adjustments to the support, they had to let go of the book, which neutralized this positive aspect.

Regarding the support adjustments, evaluators stated that the adjustments were very erratic and lacking in terms of malleability.

7.6 Hands Free + Book Support

The Figure 26 shows the results obtained in the Hands Free + Book Support Scenario.

Calculating the average SUS Score, this scenario achieved a 74.72 score. Being that six evaluators individuals scores did not achieved a score rating above 70.

The scores given by evaluators 1, 2, 7, 8, 10 and 18 resulted in the following SUS Scores: 62.5, 47.5, 50, 52.5, 42.5 and 47.5, respectively. The main complaints about this scenario is that the book support makes it difficult to turn the book pages, as they are hold by two clips in both sides.

As observed in the tests, most evaluators had to drop the smartphone to adjust the angle of the support and turn the pages, making the tasks more challenging and less intuitive.

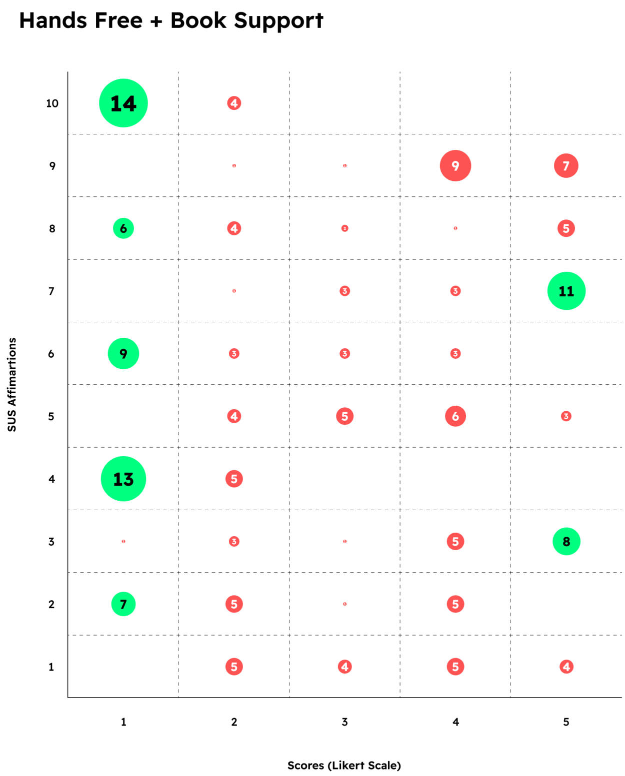


Figure 26. Hands Free + Book Support scores.

7.7 Fixed Support + Book Support

The Figure 27 shows the results obtained in the Fixed Support + Book Support Scenario.

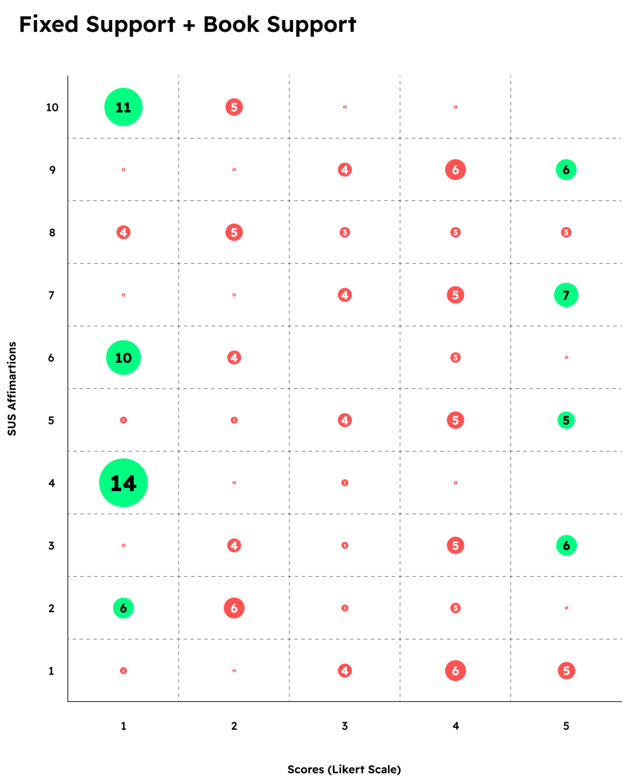


Figure 27. Fixed Support + Book Support scores.

The average SUS Score for this scenario was 71.11. Among the 10 affirmations of the SUS form, only three of them achieved a high level of consensus between the evaluators. The affirmations related to the complexity of the scenario ("I think I would need the help of a person with technical knowledge to use the system" and "I had to learn a lot of things before I could use the system") received low ratings by the majority of the evaluators, indicating that the problems about this scenario are not related to the complexity of it.

In order to understand what are the problems of this scenario we read the reports given by the evaluators which individual SUS Score did not achieve the desired 70 score. The scores given by evaluators 1, 5, 6, 9, 10, 11, 16 and 17 reached the following results: 60, 52.5, 62.5, 50, 27.5, 30, 42.5 and 62.5, respectively.

Users reported significant delays when adjusting both supports, noting that these supports also restrict interaction capability.

7.8 Articulated Support + Book Support

The Figure 28 shows the results obtained in the Articulated Support + Book Support Scenario.

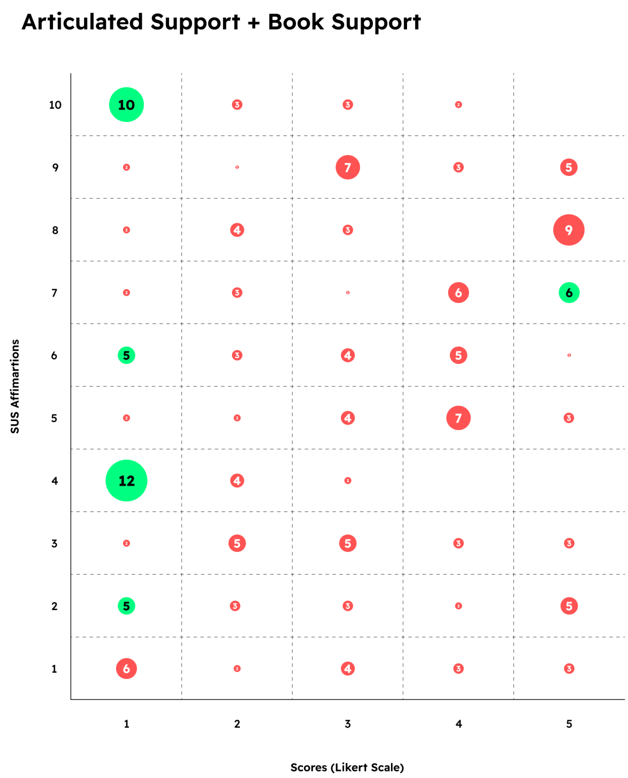


Figure 28. Articulated Support + Book Support scores.

This scenario achieved a 59.31 score, the lowest of all of them. Achieving the lowest level of consensus among the evaluators and only six of the evaluators reached an SUS Score above 70.

The main problems reported were related to the instability of the smartphone articulated support, and in some cases the evaluators had to hold the support to increase the stability. Another problem reported was the necessity of recurrent adjustments to the angle of the supports to better track the illustrations which took a lot of time.

8 Discussions and Lessons Learned

In the following Subsections we will present some discussions and lessons learned through the tests we carried out.

8.1 Regarding the Tests of AR SDKs

Among the analyzed SDKs, Vuforia proved to be the most competent for the study, showing promising results from the initial test. However, some issues can be improved and will be discussed in the next sections.

First, the illustration style used in the book chosen for this case study only partially satisfies the characteristics mentioned in the documentation of good practices of the SDK. Images must contain a high number of attributes to improve the detection and tracking capabilities of the tool. Illustrations with rounded edges do not allow attribute selection, as these rely on sharp details (90-degree angles). However, technology should allow the manifestation of different types of art.

8.1.1 Relevance of the Pages Not Classified as Acceptable

Pages classified as unacceptable are mostly pages without illustrations paired with another page, considering a reading experience with the book open. Just five single pages do not match a page rated acceptable. One of these pages is the back cover of the book. Refraining from discarding this page does not impact the reading experience, as its content is elsewhere in the book.

The four pages that detail the functioning of the Turing Machine have text with a very similar pattern and pair with each other. The content of these pages is relevant, as they try to explain in detail how the tape modifies each movement of the Turing Machine. However, this feature was probably used because inert material such as a illustrated book was limited. From the page that precedes this explanation, which consists of the page that contains the Turing Machine and the initial state of the tape, it would be possible to include an animation with the operation of the Machine and the modifications in the tape using AR resources. Such animation could contemplate all the content of the pages classified as unacceptable, reducing the impacts caused by the reading process.

The content of the other pages classified as unacceptable could be presented based on the recognition of the page with which pairing, considering a reading experience with the book open. For instance, using AR resources, feedback can be given from a page categorized as suitable for reading aloud, indicating that the content on the next page is included in the current reading material.

Other possibilities in book design could increase the number of acceptable pages. One of them would be to change the alignment of the text. However, the left alignment includes people with dyslexia; changing this format could harm this audience. The same would happen if there were a change in the font used in the text. The font used is sans serif, and this means fewer recognition points. However, a serif font would compromise the reading experience for a significant audience. So all these proposed design changes must be further analyzed to ensure no loss of functionality for people with special needs or accessibility requirements.

Finally, including new elements in the illustration could increase the number of acceptable pages. However, this strategy should be carefully analyzed with the illustrator to avoid impacting her art style. All these alternatives would generate financial impacts, considering a new printing.

8.1.2 Algorithms Should Adapt to Human Art Style

From the tests carried out with the images in the book, we portray how we can improve the images to have more extraordinary adaptability to AR. For instance, adding new elements to images or creating different image patterns.

However, while we strive to improve image recognition and tracking [Kim *et al.*, 2018], we cannot limit illustrators. The algorithms must adapt to the illustrations and not vice versa. Once the artists limit their thoughts, we restrict their authenticity and create a pattern of illustrations for books in AR.

The project's objective is, based on the results obtained in the study of the tests, to direct illustrators who wish to adapt their books later to guide and not limit their work. One of the challenges identified by this work is the study of other algorithms for AR recognition from artistic patterns and styles, which provides the same benefits for books that do not contain images.

8.1.3 Adaptation of Published Manuscripts

During the tests, we experimented with the use of QR Codes and evaluated the improvement in the performance of images that obtained scores from 0 to 2 stars. An alternative for books already sold and available to readers, although we know that it is a complicated solution to implement due to cost, acceptance, and other issues.

These findings suggest the necessity of what we term an "Extended Reality First" approach, in which the design of the physical book and its augmented components are conceived simultaneously. This perspective aligns with user-centered and interaction-driven design principles, which emphasize the importance of integrating interaction requirements from the early stages of product development. By considering augmented interaction during the conceptual phase of the book, authors and designer may reduce usability friction and structural inconsistencies between printed and digital content.

8.1.4 Devices for Interaction with AR Books

Tablets and smartphones are the most accessible mobile equipment for people to use AR applications, given the popularization of devices and their computational processing power. The results presented in Section 6.1 demonstrated that image recognition using these devices in the hands was inadequate and that the average recognition time is high.

In some cases, images were not even recognized for tracking by the SDK. There are two reasons for this fact. The first is that holding a mobile and a book simultaneously by hand is not ergonomically comfortable, as people generally use both hands to hold the book. Therefore, when using mobile for AR interaction with books, one must consider that the book is on some furniture, such as a table. The second reason is related to the instability generated by the device, which is usually in motion in the hands. Wearable computer solutions for AR are well known [Cheng *et al.*, 2021; Itoh *et al.*, 2021] and can be an alternative for these problems. However, a feasibility study should interact with books.

8.1.5 Challenges for Inclusion of Visual Impaired People

Virtual and AR technologies are the concept of extended reality. Extending reality provides additional triggers through virtual mechanisms, leading people to understand the mixture of real and virtual. For instance, removing people from the real world and inserting them into a completely virtual world. For people without visual impairment, vision and hearing are the most explored impulses to experience the concept of extension of reality. Since only vision is the sense that we use around 80% of the time when interacting with the real world [Alfailakawi, 2016]. For example, vision may be the most active sense since reading this work in a quiet environment, with hearing inactive or not processing the surrounding noises.

On the concept of reality, we will realize that, for people with visual impairments, images are not part of what reality is for them. However, touch, hearing, smell, and taste are. So, the premise is to work with virtual stimulation to these other senses. For example, a person sits in a chair blindfolded. There is a fan directed towards the person and ambient sound with the noise of a roller coaster car. As the sound of the cart on the rails demonstrates a better speed, the force of the wind generated by the ventilator intensifies. The sound (spatialization sound) and the wind made with a fan (tactile sense) can make people believe they are on a roller coaster ride even without seeing their surroundings. If this is an accurate impression, in practice, the person will be experiencing a virtual reality via touch and hearing.

8.1.6 Augmented Reality application for users gestures evaluation

The development of the application for defining user tests brought some challenges. One related to how the user interacts with the application and the physical book. During our tests, it was possible to observe that holding the mobile with one hand and holding the book with the other hand limited how we could interact with the application, making it harder to make gestures that require multiple touches on the screen. Although the interaction interface on surfaces is in 2D, where we manipulate the x and y axes, in practice, the interaction of objects can occur in a three-dimensional environment since virtual objects interact with the real world from a three-dimensional perspective. Some gestures developed for future studies were limited to a first-person perspective. This refers to the case of the slingshot event. Figure 11 shows the implemented example.

Additionally, the application users might choose not to execute actions through the mobile screen interface. For example, the rotation function can be executed using gestures on the screen surface, and the person can rotate the device around the book. There is also the option to rotate the book since AR applications allow these manipulations. This is one of the behaviors that should be observed in future tests.

Our experience raises important questions: Why is the combination of AR and books yet to be widely explored? Despite its potential to create more immersive experiences and promote reading among children and young people, there are still challenges to overcome in utilizing this technology effectively.

8.2 Regarding the Usability Evaluation

The following subsections describe some problems we noticed during the evaluation which we have not considered in advance.

8.2.1 Evaluation Duration

The biggest problem we faced during the study was related to the length of the evaluation. During the preparation phase, we anticipated that each of the six scenarios would last an average of 10 minutes. We were concerned that the process would become tiring and that volunteer fatigue could interfere with the results.

In an attempt to get around this problem, at the beginning of each evaluation, we randomly selected an order for the tests, that is, instead of always starting with the same scenario. In that way, it is possible to verify in the analysis of the results whether, over time, the volunteer's fatigue affected the perception of the different cases.

Another point to be raised is related to the time we allowed the volunteers to familiarize themselves with the application. Since this is a technology that has not yet been explored much, we noticed that some of the volunteers felt somewhat dazzled by the possibilities of practical applications of the technology. Thus, consuming a good part of the time foreseen for the tests, a possibility that was thought of later was to define a time limit for the "training" of the volunteers and to explain orally and in detail the functionalities of the application.

The point raised above is directly related to our concern about potential volunteer fatigue. We cannot allow the evaluation to go on for an indefinite period of time and allow fatigue to compromise the results obtained.

8.2.2 Materials and Infrastructure

The entire evaluation process took place with only one observer (which in this case has only a few academic experiences in evaluations involving humans, so the lack of experience must be considered) and one volunteer at a time, where the observer was responsible for receiving the volunteer, explaining the reason and objective of the experiment, as well as the TCLE, collecting socio-demographic information, and taking notes on all six different scenarios.

The way the procedure was carried out meant that some information could have gone unnoticed by the observer, since some volunteers made many comments in a short space of time, which made it difficult to take notes. Therefore, one way to get around this problem would be to use cameras or sound recorders to capture all the information given by the volunteer. As well as the presence of more observers on site to not only assist in the process of explaining the experiment, but also to speed up the entire process of filling out basic data, such as the TCLE form and the collection of socio-demographic data.

8.2.3 Volunteers' Characteristics

After the evaluation period, we observed that, in order to carry out a more robust analysis of the results obtained, we could have included in the socio-demographic data form information relevant to the proposed study. Data such as the volunteer's dominant hand, the use of glasses or contact lenses, and other aspects may interfere in some way with the evaluation responses.

It is also important to highlight that the evaluation conducted had the participation of many volunteers belonging to the academic community of the Universidade Federal de Ouro Preto. Considering this, a broader sample and in different contexts (not only university students, different locations, etc.) can bring more insights to the research.

And although the primary target audience of our study of AR-enhanced books are K-12 students, the initial evaluation was intentionally conducted with adult participants enrolled in higher education. This decision aimed to reduce confounding variables related to reading proficiency, task comprehension, and basic interaction skills, allowing the study to focus specifically on usability aspects of gesture-based interaction. By selecting participants with academic experience and familiarity with printed materials, we managed to isolate interaction-related challenges rather than difficulties associated with literacy or cognitive development stages.

Importantly, the selected participants still share a key characteristic with the intended audience: they are active students engaged in learning environments. Nevertheless, we acknowledge that motor skills, attention span, and cognitive processing differ between adults and K-12 students, which may influence gesture performance and engagement levels.

Another limitation concerns the composition of the participant sample. The distribution between men and women was not balanced, which may have influenced some perceptions of usability and technology acceptance. Future studies should seek a more balanced gender distribution.

For this reason, this study should be seen as an exploratory test. The evaluation protocol is planned to be replicated in the future with the target audience in order to validate and refine our findings.

8.2.4 Pilot Tests

All the problems reported in the previous subsections could have been predicted by carrying out more pilot (or experimental) tests, as this is a way to identify errors and improvements in the experiment even before the final evaluation is applied.

The point to be considered is the difficulty of carrying out such pilot tests without employing the participation of potential volunteers for the final experiment. Since one of the problems faced is attracting people and organizing a schedule capable of meeting the needs of both parties involved.

8.2.5 Change of Habits

Among the six scenarios evaluated, the one which achieved the highest SUS Score was the Hands Free Scenario, which participants held the book with one hand while interacting with the mobile AR application using the other. Contrary to our initial hypothesis, which assumed this configuration would impose great physical and cognitive constraints.

This can be related to cognitive familiarity. Most people are accustomed to operate smartphones with one hand while the other hand provides contextual or physical support. Alternative scenarios tested in this study required interaction patterns that deviate from common smartphone usage habits, which may have introduced discomfort rather than intrinsic usability limitations.

The Fixed Support scenario received the second-highest SUS Score, indicating a good level of usability. The question that remains is: if the evaluators had more time to become familiar with this mode of interaction, would it be possible to achieve an even higher score?

9 Conclusions

In this article, we evaluate and discuss the use of AR by analyzing a book not designed for this proposal. We chose the juvenile's book "Alan Turing: His Machines and His Secrets" as a case study. Through a qualitative analysis, we determined three SDKs for the study and selected Vuforia and ARCore. Vuforia proved to be the most efficient and accessible software for classifying the book's illustrations' quality.

The tests identify patterns of success and failure and which adaptations can be made to obtain more excellent compatibility with AR. Consequently, we approach ways to guide future illustrators who wish to adapt their books without losing their authenticity and freedom of expression. Finally, after a series of tests classifying the 63 images in the book, it was identified that 71% accepted using the hands to hold the cell phone and 76% with cell phone support.

Also, the usability evaluation carried out, explored in a practical way the main problems faced by users of AR technologies. And, despite the specific context of the work (use of AR combined with physical books), the lessons learned here are generally useful for evaluations with people.

Based on our experience, we recommend good experimental planning so that possible errors can be avoided, which could invalidate the results obtained. The main lesson we learned was the need to carry out more experimental tests, since these tests can help identify potential areas for improvement and refine the process in general.

Future work may investigate additional image enhancement techniques to improve page detection performance, evaluate the proposed approach using books with different illustration styles and graphical characteristics, and explore methods to reduce the cost and effort required to adapt existing printed books for AR applications. These investigations may contribute to improving both the scalability and practical adoption of AR-enhanced books.

Overall, we addressed the main difficulties in conducting an evaluation involving people and the errors made in the study. Thus, highlighting the need for studies that offer guidelines capable of assisting in the planning of studies of this type. Improving the study field and facilitating the understanding of the importance of the topic.

Declarations

Acknowledgements

The content of this work was written by the authors. The authors used AI tools only for grammatical proof.

Funding

This study was funded by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001, the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) finance code 306101/2021-1, and the Universidade Federal de Ouro Preto (UFOP).

Authors' Contributions

BA contributed to conducting both SDK tests and user evaluation, collecting and analyzing the data, writing and reviewing the article. SD and SB contributed to the methodology of the work and reviewing the article. PF, JG, SB and NV contributed by reviewing the article content and text review. All authors read and approved this article's version.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

The complete anonymized dataset and supporting research materials associated with this study are publicly available at Zenodo (https://doi.org/10.5281/zenodo.18836927). The repository includes: sociodemographic data (anonymized), randomized test order, individual SUS responses for all six scenarios, calculated SUS scores, TAM results.

References

Abazi-Bexheti, L., Kadriu, A., and Apostolova, M. (2022). Research on vr/ar integration in education. In *2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO)*, pages 563–567. DOI: https://doi.org/10.23919/MIPRO55190.2022.9803398.

Adachi, B., Bim, S., and Delabrida, S. (2025). Lessons learned in human evaluation: A case study of augmented reality in books. In *Proceedings of the ACM International Conference on Interactive Media Experiences Workshops*, pages 136–139, Porto Alegre, RS, Brasil. SBC. DOI: https://doi.org/10.5753/imxw.2025.8905.

Adachi, B. H., Gonzaga, J. C., Fernandes, P. C., Silva, S. E. D., Bim, S. A., and Boss, S. L. B. (2024). Augmented reality in books: An evaluation of alan turing book. In *Proceedings of the XXII Brazilian Symposium on Human Factors in Computing Systems*, IHC '23, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/3638067.3638124.

Al-Ali, H., Bazzaza, M. W., Zemerly, M. J., and Ng, J. W. P. (2016). Myvision air: An augmented interactive reality book mobile application. In *2016 IEEE Global Engineering Education Conference (EDUCON)*, pages 741–745. DOI: https://doi.org/10.1109/EDUCON.2016.7474634.

Alfailakawi, A. (2016). The effects of visual training on vision functions and shooting perfomance level among young handball players. *Ovidius University Annals, Series Physical Education & Sport/Science, Movement & Health*, 16(1).

Apple (2026). Detecting images in an ar experience. `https://developer.apple.com/documentation/arkit`. Accessed on 09 July 2026.

Bittencourt, P. A. S. and Albino, J. P. (2017). O uso das tecnologias digitais na educação do século xxi. *Revista Ibero-Americana de Estudos em Educação*, page 205–214. DOI: https://doi.org/10.21723/riaee.v12.n1.9433.

Boss, S. L. B. and Bim, S. A. (2022). *Alan Turing: suas máquinas e seus segredos*. Editora Blucher, São Paulo, SP, Brasil.

BRASIL (2018). Base Nacional Comum Curricular. `http://basenacionalcomum.mec.gov.br`. Accessed on 09 July 2026.

BRASIL (2022). Base Nacional Comum Curricular: Computação – Complemento à BNCC. `https://basenacionalcomum.mec.gov.br/images//historico/anexo_parecer_cneceb_n_2_2022_bncc_computacao.pdf`. Anexo ao Parecer CNE/CEB nº 2/2022. Accessed on 09 July 2026.

Brooke, J. (1995). Sus: A quick and dirty usability scale. *Usability Eval. Ind.*, 189.

Chen, Y., Wang, Q., Chen, H., Song, X., Tang, H., and Tian, M. (2019). An overview of augmented reality technology. *Journal of Physics: Conference Series*, 1237:022082. DOI: https://doi.org/10.1088/1742-6596/1237/2/022082.

Cheng, D., Wang, Q., Liu, Y., Chen, H., Ni, D., Wang, X., Yao, C., Hou, Q., Hou, W., Luo, G., and Wang, Y. (2021). Design and manufacture ar head-mounted displays: A review and outlook. *Light: Advanced Manufacturing*, 2(3):350–369. DOI: https://doi.org/10.37188/lam.2021.024.

Cheng, K.-H. (2017). Reading an augmented reality book: An exploration of learners' cognitive load, motivation, and attitudes. *Australasian Journal of Educational Technology*, 33(4). DOI: https://doi.org/10.14742/ajet.2820.

Cheng, K.-H. (2019). Parents' user experiences of augmented reality book reading: perceptions, expectations, and intentions. *Educational Technology Research and Development*, 67(2):303–315. DOI: https://doi.org/10.1007/s11423-018-9611-0.

Cumbo, B. and Selwyn, N. (2021). Using participatory design approaches in educational research. *International Journal of Research Method in Education*, 45:1–13. DOI: https://doi.org/10.1080/1743727X.2021.1902981.

Danaei, D., Jamali, H. R., Mansourian, Y., and Rastegarpour, H. (2020). Comparing reading comprehension between children reading augmented reality and print storybooks. *Computers & Education*, 153:103900. DOI: https://doi.org/10.1016/j.compedu.2020.103900.

De Barros, A. F. (2019). O uso das tecnologias na educação como ferramentas de aprendizado. *Revista Científica Semana Acadêmica*, pages 01–13.

Dede, C. (2009). Immersive interfaces for engagement and learning. *Science*, 323(5910):66–69. DOI: https://doi.org/10.1126/science.1167311.

Del-Masso, M. C. S., Cotta, M. A. d. C., and Santos, M. A. P. (2014). Ética em pesquisa científica: conceitos e finalidades. *São Paulo: Núcleo de Educação a Distância da Unesp*.

Demitriadou, E., Stavroulia, K.-E., and Lanitis, A. (2019). Comparative evaluation of virtual and augmented reality for teaching mathematics in primary education. *Education and Information Technologies*, 25(1):381–401. DOI: https://doi.org/10.1007/s10639-019-09973-5.

Dias, B., Keller, B., and Delabrida, S. (2019). Evaluation of augmented reality sdks for classroom teaching. In *Proceedings of the 18th Brazilian Symposium on Human Factors in Computing Systems*, IHC '19, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/3357155.3358447.

Dong, C. and Si, Z. (2018). *The Research and Application of Augmented Reality in 3D Interactive Books for Children*, pages 293–299. DOI: https://doi.org/10.1007/978-981-10-7629-9_35.

Dong, Z., Zhang, J., Bai, X., Clark, A., Lindeman, R. W., He, W., and Piumsomboon, T. (2022). Touch-move-release: Studies of surface and motion gestures for mobile augmented reality. *Frontiers in Virtual Reality*, Volume 3 - 2022. DOI: https://doi.org/10.3389/frvir.2022.927258.

Freitas, S. (2018). Are games effective learning tools? a review of educational games. *Educational Technology and Society*, 21:74–84.

Friesen, S. and Jacobsen, M. (2021). Collaborative design of professional graduate programs in education. *International Journal of Designs for Learning*, 12(1):64–76. DOI: https://doi.org/10.14434/ijdl.v12i1.25778.

GitHub (2025). What is an sdk? `https://github.com/resources/articles/what-is-an-sdk?locale=pt-BR`. Accessed on 11 July 2026.

Google (2026). Add dimension to images. `https://developers.google.com/ar/develop/augmented-images`. Accessed on 09 July 2026.

Interaction Design Foundation (2016). User centered design (ucd). `https://www.interaction-design.org/literature/topics/user-centered-design`. Accessed on 11 July 2026.

Itoh, Y., Langlotz, T., Sutton, J., and Plopski, A. (2021). Towards indistinguishable augmented reality: A survey on optical see-through head-mounted displays. *ACM Comput. Surv.*, 54(6). DOI: https://doi.org/10.1145/3453157.

Kim, K., Billinghurst, M., Bruder, G., Duh, H. B.-L., and Welch, G. F. (2018). Revisiting trends in augmented reality research: A review of the 2nd decade of ismar (2008–2017). *IEEE Transactions on Visualization and Computer Graphics*, 24(11):2947–2962. DOI: https://doi.org/10.1109/tvcg.2018.2868591.

Kryeziu, S. A., Avdiu, T. A., and Avdiu, A. (2021). Examining the teachers, administrators and parents' view on drawbacks of technology use in education. *Ilkogretim Online*, 20(2). DOI: https://doi.org/10.17762/pae.v58i2.3713.

Lai, A. S., Wong, C. Y., and Lo, O. C. (2015). Applying augmented reality technology to book publication business. In *2015 IEEE 12th International Conference on e-Business Engineering*, pages 281–286. DOI: https://doi.org/10.1109/ICEBE.2015.55.

Ma, Q. and Liu, L. (2005). *The Technology Acceptance Model*. DOI: https://doi.org/10.4018/9781591404743.ch006.ch000.

Maciel, C., Bim, S. A., and da Silva Figueiredo, K. (2018). Digital girls program: disseminating computer science to girls in brazil. New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/3195570.3195574.

Mann, S., Havens, J. C., Iorio, J., Yuan, Y., and Furness, T. (2018). All reality: values, taxonomy, and continuum, for virtual, augmented, extended/mixed (x), mediated (x, y), and multimediated reality/intelligence. In *Presented at the AWE 2018*. DOI: https://doi.org/10.48550/arXiv.1804.08386.

Mazon, J., de Castro Pasquini, G., da Rosa Pimentel, A. C., Souza Capistrano, V., Lucas de Oliveira Ramalho, J., Marino Costa, A., Pozzebon, E., and Bolan Frigo, L. (2022). Curso rafec - realidade aumentada facilitando o ensino das ciências: Uma nova possibilidade ao professor. *RECIMA21 - Revista Científica Multidisciplinar*, 3(8):e381797. DOI: https://doi.org/10.47820/recima21.v3i8.1797.

Ribeiro, L., Foss, L., Cavalheiro, S. A. D. C., Kniphoff da Cruz, M. E. J., and Soares de França, R. (2023). The brazilian school computing standard. SIGCSE 2023, page 53–58, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/3545945.3569863.

Skarbez, R., Smith, M., and Whitton, M. C. (2021). Revisiting milgram and kishino's reality-virtuality continuum. *Frontiers in Virtual Reality*, Volume 2 - 2021. DOI: https://doi.org/10.3389/frvir.2021.647997.

Søegaard, M. (2024). System usability scale for datadriven ux. `https://www.interaction-design.org/literature/article/system-usability-scale`. Accessed on 11 July 2025.

Unsplash (2026). Unsplash. `https://unsplash.com/pt-br`. Accessed on 09 July 2026.

Vuforia (2026). Best practices for designing and developing image-based targets. `https://developer.vuforia.com/library/vuforia-engine/images-and-objects/image-targets/image-targets/`. Accessed on 09 July 2026.

Zambiasi, P. and Pozzebon, E. (2025). Provadores virtuais de realidade aumentada: Diferenças geracionais na aceitação da tecnologia: Um estudo experimental no contexto brasileiro de e-commerce de moda. *Revista de Geopolítica*, 16(5):e811. DOI: https://doi.org/10.56238/revgeov16n5-043.