

RESEARCH

Open Access



# Perspectives on how to evaluate augmented reality technology tools for education: a systematic review

Manoela M. O. da Silva<sup>1\*</sup> , João Marcelo X. N. Teixeira<sup>1,2</sup>, Patrícia S. Cavalcante<sup>3</sup> and Veronica Teichrieb<sup>1</sup>

## Abstract

Education has benefited from augmented reality's (AR) potential to promote interactive experiences both inside and outside the classroom. A systematic review was conducted on how AR's impact in the learning process has been evaluated. We selected papers from 2009 to 2017 in three databases, IEEE, ACM, and Science Direct, using an open-source crawler, and in one Brazilian Conference, SBIE. We followed the PRISMA protocol. Forty-five works were selected and used to extract data for our research. They were also analyzed according to quantitative and qualitative criteria. The results from all the papers are available in an online database. Results evidenced an increase in the number of papers evaluating the AR's impact in education. They also showed that AR has been applied in different areas and contexts. Most papers reported positive outcomes as a result of AR insertion. However, most studies lacked the involvement of the teacher and the use of multiple metrics to evaluate educational gains.

**Keywords:** Augmented reality, Evaluation, Educational systems

## Introduction

Augmented reality (AR) is a technology that consists of adding virtual elements to a real scene coherently so that ideal users cannot differentiate them from the real scene [3]. Although all fields of knowledge can potentially take advantage from AR, Tori et al. [72] argue that education will be particularly modified by its introduction. The coexistence of virtual and real environments allows learners to experience phenomena that otherwise would be impossible in the real world. This allows learners to visualize complex spatial relationships and abstract concepts and, therefore, develop important abilities that cannot be evolved in other technology learning environments [78].

It has been long since AR's potential in education has been investigated. According to Kostaras et al. [41], AR can aid learning and make the overall process more interesting and pleasant. In a rapidly changing society as ours where there is a great amount of information available, it is of major importance to know how to locate information

and use it efficiently. AR dramatically shifts the location and timing of education and training [46].

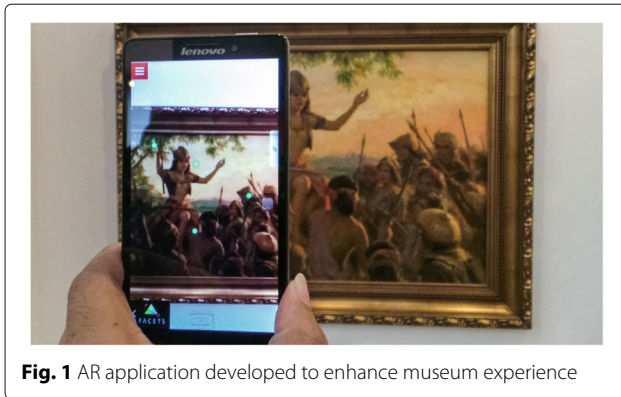
Billinghurst and Duenser [5] explain that unlike other computer interfaces that draw users away from the real world and onto the screen, AR interfaces enhance the real world experience as shown in Fig. 1, which presents an AR application designed to create new museum experiences [2]. Billinghurst and Duenser [5] also highlight some reasons why AR educational experiences are different: (a) support of seamless interaction between real and virtual environments, (b) use of a tangible interface metaphor for object manipulation, and (c) ability to transition smoothly between reality and virtuality.

Although AR has been studied for over 40 years, only in the last decade it began to be formally evaluated [23, 24, 68]. One of the reasons why it took so long to have user evaluations may be a lack of knowledge on how to properly evaluate AR experiences and design experiments [24]. Dünser et al. [24] claim that there seems to be a lack of understanding regarding the need of doing studies and the right motivation for carrying them. If user evaluations are conducted out of incorrect motivation or if empirical methods are not properly applied, the findings are of limited value or can even be misleading.

\*Correspondence: [mmos@cin.ufpe.br](mailto:mmos@cin.ufpe.br)

<sup>1</sup>Voxar Labs, Centro de Informática, Universidade Federal de Pernambuco, Recife, Brazil

Full list of author information is available at the end of the article



**Fig. 1** AR application developed to enhance museum experience

Until that time, the amount of AR systems formally evaluated was rather small [23]. Swan and Gabbard [68] and Dünser et al. [24] have found that only around 8% of published AR research papers included formal evaluations. According to Dünser and Billingham [22], one reason for this small percentage may be the lack of suitable methods for evaluating AR interfaces. Researchers in non-conventional interface fields such as virtual reality (VR) or AR cannot rely solely on design guidelines for traditional user interfaces since new interfaces afford new forms of interactions [22]. Since then, more works address some form of user evaluation [6].

When dealing with educational AR systems, it is also important to evaluate the impact of learning applications and the feasibility of incorporating them into the classrooms. Many factors are involved in this process varying from cost to staff's acceptance. Evaluation of technology is an important step in design instruction, which is the process by which learning products and experiences are designed, developed, and delivered<sup>1</sup>. Also, it is necessary to evaluate it properly so practitioners are more confident in its positive effects. It is also relevant to consider the points of view of both teachers and learners since they might differ.

In the last decade, a few papers have been published evaluating educational aspects of AR applications used for education. For instance, Balog and Pribeanu [4] have shown the same aspect can be valued differently by both teachers and learners. One survey reviewed applications intended to complement traditional curriculum materials for K–12 [65]. It performed a qualitative analysis on the design aspects and evaluation for AR Learning Environments (ARLES). Its focus was to investigate ARLES designed for kindergarten and primary and/or secondary school, as well as to explore learning theories as basis for effective learning experiences. They found out that there are three inherent AR affordances to educational settings: real-world annotation, contextual visualization, and vision-haptic visualization [65]. These affordances were supported by existing theories. Authors discovered

that aside from the performance of students in pre- and post-tests, other aspects of the learning experience such as motivation and satisfaction were usually observed.

However, it can be noted that the aforementioned paper focuses only on K–12 education. Our paper will focus on different target groups of the AR applications evaluated.

As this research area matures and the use of AR in education grows, it is important to analyze its impact appropriately to have relevant and valid feedback for the stakeholders involved in the process. Thus, this paper presents a systematic review on how studies have been evaluating AR in education.

The contributions of this paper are:

- The use of a robust research methodology to collect and analyze papers that perform educational evaluation of AR educational applications (“[Methodology](#)” section)
- A classification and discussion of studies that evaluate educational aspects of such AR systems (“[Results and discussion](#)” section)
- Guidelines to evaluate educational aspects of AR applications (“[Guidelines for educational evaluation](#)” section)

## Methodology

Considering the complexity of the educational field, such as different learning needs and times, to name a few, and its implications for technology acceptance and use, a systematic review was conducted to investigate how researchers are evaluating their AR systems. This review followed the PRISMA protocol [57] as shown in Fig. 2.

## Research questions

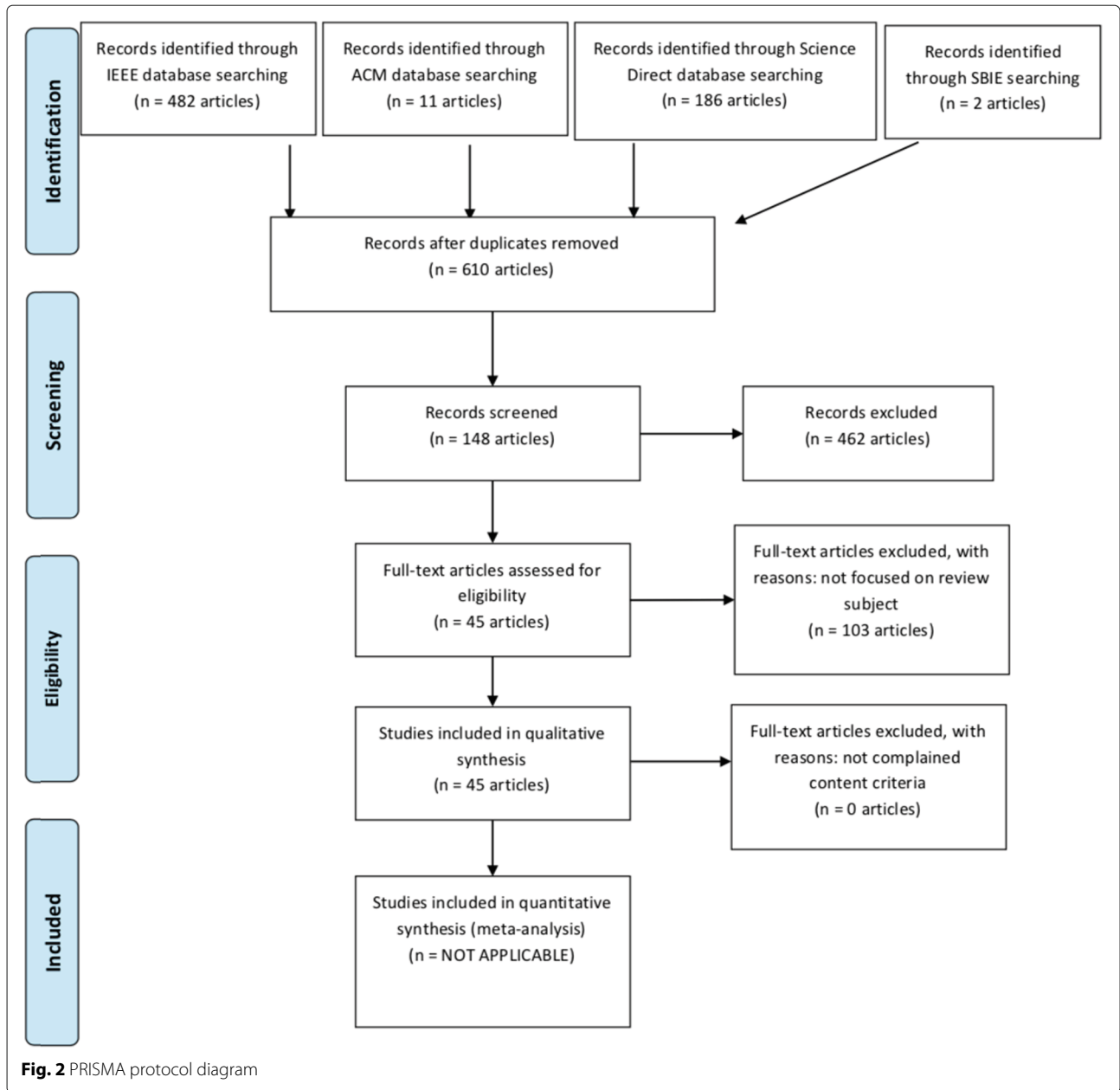
Our main question was “how do researchers evaluate AR-based educational technology?”. To guide data extraction, analysis, and synthesis, sub-questions were formulated as listed below. The questions are divided into three categories: descriptive, classificatory, and relation and effect.

### *Descriptive questions:*

1. What is the evolution in number and type of research from 2009 to 2017?
2. What institutions are most involved in performing this type of research?

### *Classificatory:*

3. What are the different designs (methodologies) used in these studies?
4. What are the target populations used in these studies?
5. What are the constructs being analyzed?
6. What are the domains of the different applications tested?
7. What types of research questions are investigated?



8. What are the types of AR technology used?
9. What is the problem being analyzed?
10. Is the application based on any educational theory?
11. What technologies AR is combined with?
12. How was the involvement of teachers in the evaluation process?
13. Did the study use multiple metrics (both quantitative and qualitative)?
14. Did the study use multiple metrics for educational evaluation purposes?

*Relation and effect:*

15. What is the kind of impact of the tool analyzed?

**Systematic review procedure**

The first step was to establish the search string for paper selection. The search string was created based on our research questions. The terms were defined along with synonyms found in the literature as shown in Table 1.

Then, the databases for the search were defined. Papers were searched automatically in three databases: ACM, IEEE Xplore, and Science Direct. Also, papers were searched in the main Brazilian Conference related to Informatics in Education, the Brazilian Symposium on Informatics in Education, SBIE. This search was performed manually in the Google Scholar platform using our search string.

**Table 1** Search string used in the systematic review

Search string
("augmented reality") AND ("education" OR "learning")
AND ("evaluation") AND ("educational tool" OR
"educational system" OR "educational application" OR
"educational platform")

The automatic search was performed in the databases using the same open-source paper crawler software that was used by Roberto et al. [63]. This crawler enabled authors to automate the process of retrieving papers. It uses only the search string as input, and it accesses the digital libraries to search in the title, abstract, and keywords of each paper. The crawler collects the papers, eliminates duplicate versions, and creates a spreadsheet containing all the works with their title, year, source, primary affiliation, abstract, and web address.

For papers to be included in the study, they must meet the following criteria:

1. Papers published in English with more than four pages
2. Papers were only considered once (in case of repetitive papers, we considered the more complete or the most recent one)
3. Papers published from 2009 to 2017
4. Papers that explicitly mentioned their evaluation methodology
5. The papers must have at least an AR prototype working
6. The AR solution must be tested with its end users
7. The solutions presented must be applied to learning a new concept or skill
8. Papers that intended to evaluate learning aspects

First, a search was performed in the databases using the search strings. Then, in the pre-selection phase, the researchers screened the papers by reading their title, abstract, and conclusion to eliminate the ones clearly not related to the research question. Later, we applied the inclusion criteria to those papers. These papers were screened to evaluate their quality concerning quantitative and qualitative aspects. In the extraction phase, we read the papers to extract relevant data concerning the research questions.

#### Data extraction

We extracted relevant information from the selected papers as listed below. The data was organized in a spreadsheet.

- Title
- Year
- Authors
- University/research group

- Source (conference or journal)
- Methodology design
- Target population
- Application domain
- Type of research question
- Implications for practice
- Type of AR technology (tracking, display, interaction)
- What constructs does it evaluate?
- Is the application based on educational research?
- What technologies AR is combined with?
- How was the involvement of teachers in the evaluation process?
- Did the study use multiple metrics (both quantitative and qualitative)?
- Did the study use multiple metrics for educational evaluation purposes?
- What are the implications of the findings in research and practice?
- What is the impact of the tool analyzed (positive or negative)?
- Observations

#### Quality criteria evaluation

The QualSyst standard was used as a guideline for quality control [40]. This questionnaire consisted of 14 items evaluating study questions concerning design methodology, sample, outcomes, results outcomes, description, and conclusions. Some items were not scored due to their non-applicability in the study's methodology (e.g., evaluator and user blinding); in these cases, we used n/a (not applicable) in the table. Other items such as interventional and random allocation were applied only in some cases. Each item was graded as it fulfilled the requirements in three categories: total, partial, and none with assigned scores of 2, 1, or 0, respectively. The total sum was divided by the maximum possible points (e.g., 10 items  $\times$  2 points = 20 points). The final score of each paper formed a grade. In case the paper did not conduct one type of research, qualitative or quantitative, a dash was used (-) to represent this situation in the spreadsheet.

#### Threats to validity

Authors are aware of the importance of considering threats to validity in order to judge the systematic review strengths and limitations. The main issues in this type of research are related to incomplete sets of relevant papers and researcher bias regarding quality analysis.

Limitations with search string, scientific databases, and search strategy can result in an incomplete set of relevant papers. As a way to mitigate the risks, the following strategies were used: first, in order to validate the search string, the terms were discussed among the authors. The authors have a different set of skills, two of them hold Ph.D. degrees in the field of Computer Science, one has

a Ph.D. degree in Education, and one has a B.A. in Languages and is currently a Ph.D. candidate in Computer Science. All of them are teachers with experience in different educational levels, from early childhood education to post-graduate education. Second, the scientific databases that publish works from the most important conferences and journals in the area were selected, along with the papers published in the main Brazilian conference in the area. Third, the crawler uses a different approach to maximize the number of papers found. Instead of using the complete search string, eight different searches were performed using the combination of every term in both parts of the search string, which increases the number of papers collected [63].

The qualitative analysis of the papers was conducted by one of the authors. Since this may lead to a researcher bias, 15% of the papers were randomly selected to compose a set of control papers in order to increase credibility. The other authors examined the control papers to analyze them concerning their quality. The authors compared their results using Cohen's Kappa coefficient, which measures the agreement between the two classifications taking into account how much agreement would be expected to be present by chance [11]. The coefficient lies between  $-1.0$  and  $1.0$  in which  $1.0$  denotes perfect agreement,  $0.0$  indicates that any agreement is due to chance, and negative values present agreement less than chance. There is no consensus on what are good levels of agreement. Nevertheless, studies [1] mention that there is no agreement for negative values, poor agreement between  $0.00$  and  $0.20$ , fair agreement between  $0.21$  and  $0.40$ , moderate agreement between  $0.41$  and  $0.60$ , good agreement between  $0.61$  and  $0.80$ , and very good agreement for values higher than  $0.80$ . In our work, the qualitative analysis Cohen's Kappa was  $0.7969$ , which is close to very good agreement among the authors.

## Results and discussion

This section describes and discusses the results of the systematic review.

The search in the databases using the search strings returned 607 articles, and 148 papers remained after the pre-selection phase. Finally, after applying the inclusion criteria, 45 papers were eligible for this study. The results from all the papers are available in an online database, which can be collaboratively updated<sup>2</sup>.

### Quality of report

The quantitative and qualitative assessments are available at Appendixes A.1 and A.2, respectively.

### Descriptive questions

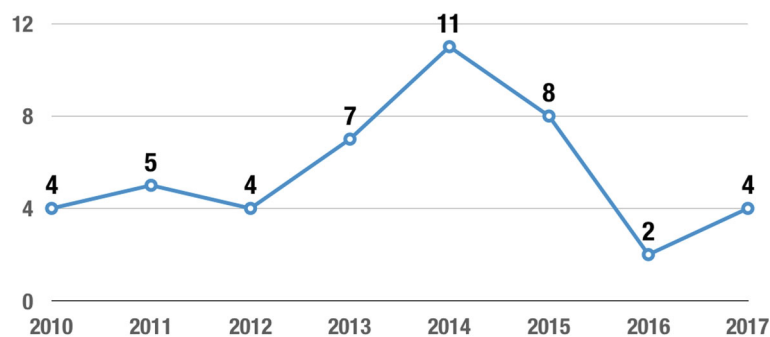
Questions 1 and 2 are in this category. Figure 3 shows that although no research was found in 2009, the research in this field is steadily growing, reaching the highest number of papers in 2014. Although the number of papers per year has decreased compared to 2014, we observe that the interest in evaluating AR for education remains.

Table 2 presents the institutions involved in the research.

Table 3 shows the venues where the studies have been published.

Figure 4 evidences that the methodology most commonly used is the experimental design, while the quasi-experimental design appeared in fourth place. The essential feature of experimental research is that the researcher deliberately controls and manipulates the conditions, which determine the events of interest [12]. Quasi-experiments are used when subjects must be allowed to choose their treatment, which is the main difference when compared to experimental designs.

Questionnaires were the second most popular method among the studies. They consist in a series of questions or prompts aimed at gathering information from subjects. The questionnaires used in the papers were designed in different ways and for varied purposes. As examples, Zhang et al. [82] used a questionnaire to investigate flow experience and Wei et al. [77] to assess creative design learning motivation. In turn, Ibáñez et al. [36] designed an open-ended questionnaire. Regardless of its structure and



**Fig. 3** Papers according to the year of publication



**Table 2** Institutions involved in the research

Institution	Papers
Federal University of Pernambuco	3
National Taiwan Normal University	3
Ramon Llull University	3
National Hsinchu University of Education	2
Universidad Carlos III de Madrid	2
Other universities with one published paper	32

aim, Cohen et al. [12] point out that an ideal questionnaire must be clear and unambiguous.

Observations appeared in seven studies while only one work reported a case study. Merriam [56] explains that observations take place where a given phenomenon naturally occurs. She points out that the skills to be a good observer must be learned; thus, training and mental preparation is important. She highlights the need to define what to observe as well as to write careful and useful field notes.

Case study, on the other hand, is an empirical inquiry that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident [81]. Merriam [56] points out that the most defining characteristic of a case study lies in delimiting the case to be studied. Thus, case study research uses purposive sampling rather than random sampling [25].

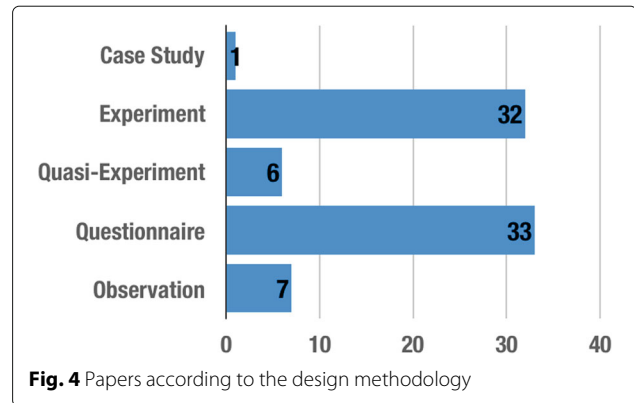
It is important to highlight that a high number of papers (32, total) reported a combination of methods or metrics. The most common combination is the experiment coupled with questionnaires. However, these multiple metrics usually not only evaluated education, but also other aspects such as motivation and satisfaction. The results for this question also evidenced a predominance of quantitative methods in the works.

Question 4 refers to the target population of the studies as seen in Fig. 5.

Figure 5 shows that the most popular target audience are undergraduate students and elementary school children.

**Table 3** Venues that published the selected papers

Venue	Papers
Computers & Education	12
International Wireless Communications and Mobile Computing Conference	2
IEEE Frontiers in Education Conference (FIE)	2
International Conference on Virtual and Augmented Reality in Education	2
Simpósio Brasileiro de Informática na Educação	2
Other venues with one published paper	25

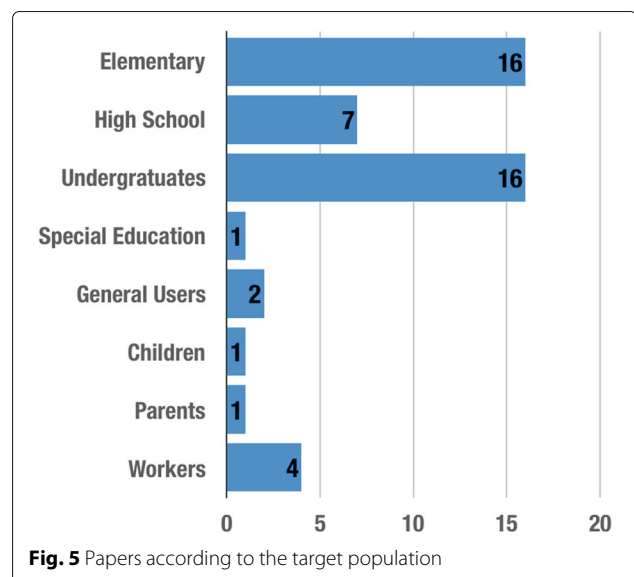


High school students appeared in seven works, thus being the third most popular audience for AR tools.

Other groups were also considered in these papers. For instance, two papers targeted general audiences of users. For instance, in Sommerauer and Müller [67], the population was an exhibition audience, which included heterogeneous genders, age groups, and educational levels. These varied target populations show that AR can expand the barriers of the school setting and achieve both formal and informal environments.

Also, parents were the audience in Cheng and Tsai [10]. This paper also targets children in both elementary and preschool. Tobar-Muñoz et al. [71] present an AR tool for children with varied ages and special needs. Finally, four papers are targeted to workers in different fields, such as engineering [8] and surgery [45]. This data evidence that AR can also be successfully used for training.

Hence, data show that although there has been a preference for undergraduate students and elementary school children, AR can be used by a variety of people, with different needs and in different contexts.



Question 5 was about the constructs evaluated in the studies as displayed in Fig. 6.

Figure 6 reveals that many studies did not evaluate solely educational aspects. Twelve works evaluated more than one aspect. The majority of the papers evaluated knowledge retention or performance.

Some applications were under development or had been recently developed; thus, usability aspects, such as users' attitudes and satisfaction, were also analyzed. Martín-Gutiérrez et al. [52] point out that the study was carried out with the beta version of the tool, which was tested with 235 students. These authors, thus, also evaluated user's satisfaction. In turn, Tarng et al. [70] investigated the attitudes of experimental group students after using the AR system. The authors explain that the questions in their study were categorized in learning contents, interface design, and applications.

Behavior and motivation were also evaluated in eight studies. Other studies evaluated constructs related to the theories they used, such as flow experience [8, 36] and dimensions of learning style [49].

Other aspects evaluated were creativity [77], teaching effects [77], and learner's opinions [69]. This variety evidences that due to the complexity of the learning environment, different aspects can be the focus of educational or learning evaluation. Depending on the focus of the studies, such as training, authors would focus on more mechanical aspects such as precise skills development and time. Conversely, studies focusing on the school environment may focus their attention on the role of the teacher, flow experience, or student's motivation to learn.

Question 6 concerns the application domains of knowledge as shown in Fig. 7.

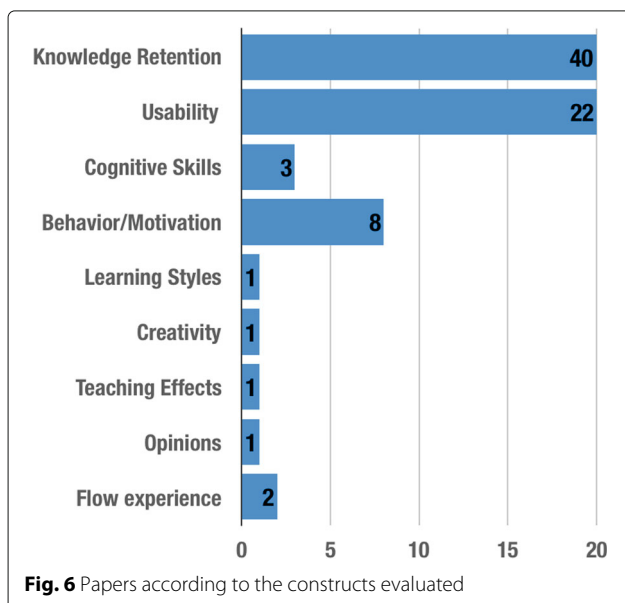


Figure 7 shows that most AR tools are related to STEM fields. STEM is an acronym that refers to the fields of science, technology, engineering, and mathematics. The second most popular domain for applications are humanities, followed by medicine and health.

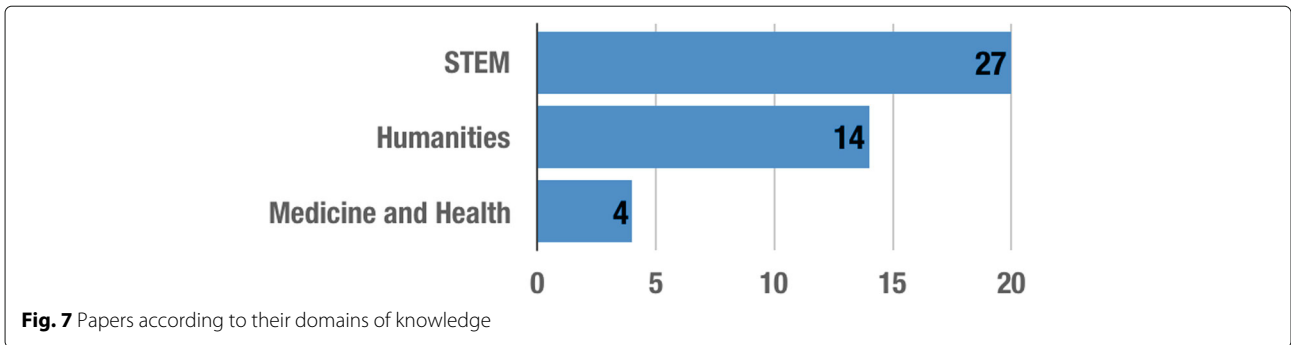
Question 7 investigated the types of research questions in the works. The questions were classified according to their types as proposed by Easterbrook et al. [25]. These authors divide research questions in two types: design questions, which are usually asked by software engineers in order to better ways to do software engineering, and knowledge questions, which are described below:

- Exploratory questions: are asked in the early stages of research when researchers are attempting to understand the phenomena, e.g., existence questions, description and classification, descriptive comparative
- Base-rate questions: are frequently asked after having a clearer understanding of the phenomena. They might be frequency and distribution questions and descriptive process
- Relationship questions: are meant to understand the relationship between two different phenomena
- Causality questions: are an attempt to explain why a relationship holds and identify its cause and effect, e.g., causality questions, causality-comparative questions and causality-comparative-interaction questions.

Figure 8 presents the types of research questions found in the papers.

Twenty-three papers asked more than one question. The chart shows that the majority of the papers asked relationship questions; those papers aimed to describe the effect of AR compared to other resources and its relationship with different aspects (e.g., academic achievement or motivation). The second most common type of question was exploratory ones, mainly descriptive comparative (present in 19 papers). Design questions were asked by two studies and causality ones by one study. This amount of exploratory questions may indicate that research in the use AR tools for education might still be in early stages, in which researchers attempt to better understand the field and the implications of such technology in education. Also, they want to understand what are the better ways to develop their tools, as evidenced in the design questions.

AR technologies used in the studies were classified according to their tracking, display, and interaction techniques. As concerns the displays used, Fig. 9 shows that screen-based and handheld were the most frequent used displays (21 and 14 papers, respectively). Screen-based displays are known for their cost-efficiency since they require off-the-shelf hardware and standard PC



equipment. They are also largely present in schools nowadays and were usually well evaluated by users.

On the other hand, the popularization and technical advancements in smartphones make handheld displays a good option for AR applications. These devices are minimally intrusive and highly mobile [83]. They enable high flexibility, as shown in Jerry and Aaron [38], in which a context-aware AR learning solution is proposed as a scaffolding platform for outdoor field learning. Tarng et al. [70] used these displays to provide situated learning.

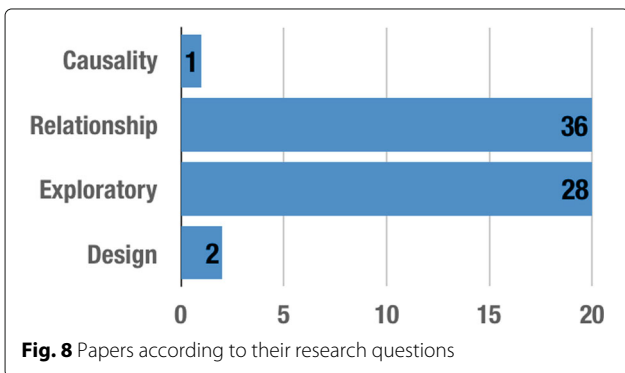
Two papers used head-attached displays. Although these displays provide a better field of view, fashion constraints are a common issue. For instance, Martín-Gutiérrez et al. [52] reported that the HMD use was not comfortable. The cables linking the glass and camera with the PC interfered with user’s movement.

Five studies presented spatial displays. Three studies did not provide enough information about the system evaluated; therefore, it was not possible to classify these systems in all three categories of AR [27, 28, 38].

As regards to tracking, 35 papers presented vision-based tracking and five papers presented sensor-based tracking as shown in Fig. 10. Two papers presented hybrid tracking.

Vision-based tracking can be divided in two categories, marker-based and markerless as illustrated in Fig. 10.

Marker-based was the most common type found (25 works). It is a very popular choice since there are many marker-based kits available for a low cost. Most papers presented positive outcomes regarding these tools. However, markers can be intrusive in the scene.



On the other hand, markerless systems do not require the use of markers. In this case, the environment itself acts as a marker. It allows guidance information to be superimposed on a real game board, for example. This type was chosen in ten studies.

Finally, as regards to interaction techniques, 19 papers presented a more traditional type of interaction using buttons, touch, or simply providing visualization of the augmented content. Their use was generally positive.

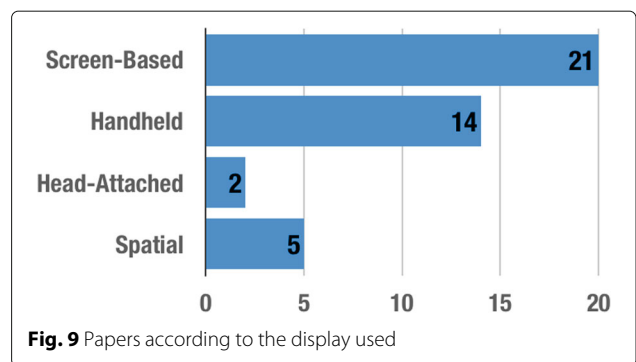
The second most common choice was tangible interaction (13 papers). These interfaces are promising as they take advantage of the familiarity of everyday objects to ease the interaction. Their use provided positive results.

Haptic interfaces were chosen in four papers. One paper presented collaborative interaction [48]. No papers chose hybrid interfaces.

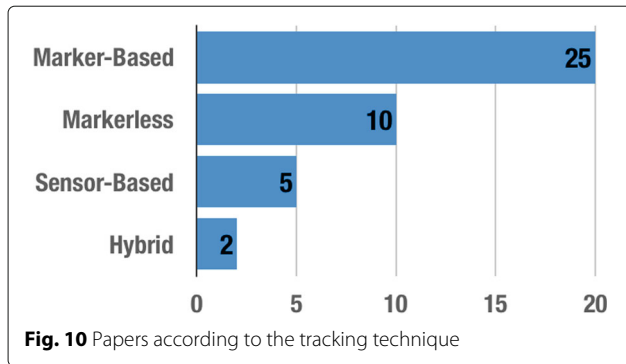
Figure 11 shows the interaction techniques used in the papers.

Question 11 investigated if the studies were based on any educational theory as presented in Table 4.

Table 4 evidences that most papers mentioned educational theories. However, 19 studies did not mention any theory. The most mentioned theories were situated learning theory and the cognitive theory of multimedia learning and cognitive load theory, mentioned in three works each. The situated learning theory emphasizes the reality of learning activities; thus, the context in which the activity naturally occurs is indispensable. AR allows real-life experiences to be enhanced with virtual content, hence expanding learning horizons.





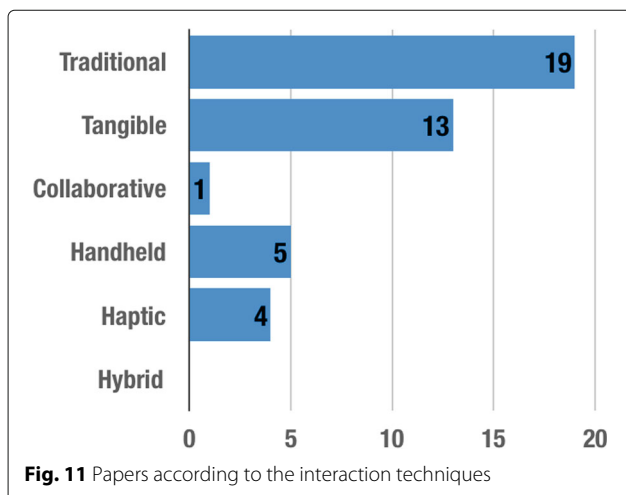


**Fig. 10** Papers according to the tracking technique

In turn, the cognitive theory of multimedia learning (CTML) states that people learn better from pictures and words rather than pictures alone. This theory is based on three assumptions: (a) people possess two channels for processing information (the auditory/verbal and visual/pictorial), (b) there is a limited amount of information each channel can process at a time, and (c) learning is an active process of selecting relevant information, organizing them into coherent mental representations, and finally integrating those representations with existing knowledge [54].

Inquiry-based learning was mentioned in two studies. As an example, Jerry and Aaron [38] mentioned this theory, which is an approach to teaching and learning that places students’ questions, ideas, and observations at the center of the learning experience [60]. Hutchings [34] adds that the process of inquiry is in the ownership of the learners; thus, inquiry-based learning is fundamentally concerned with establishing the context within which inquiry may best be stimulated and students can take charge of their learning.

Mobile learning was mentioned in two works. Mobile learning or, simply, m-learning is the didactic-pedagogical expression used to designate a new educational “paradigm” based on the use of mobile technologies [58]. Also, McGreal [55] adds that “m-learning happens in



**Fig. 11** Papers according to the interaction techniques

**Table 4** Papers for the most used education theory

Study reference	Educational theory
[50], [76], [17], [52], [51], [8], [28], [27], [64], [32], [37], [80], [62], [7], [31], [45], [18], [30], [20]	Not mentioned
[74], [70], [9]	Situated learning
[33], [67], [39]	Cognitive theory of multimedia learning and cognitive load theory
[42], [82]	Cognitive load theory
[73], [39]	Learning styles theory
[35], [38]	Mobile learning

context in which it is needed and relevant and is situated within the active cognitive processes of individual and groups of learners.” Thus, it takes advantage of the widely available mobile devices to provide access to learning anywhere and anytime, which changes many paradigms of traditional education.

The learning styles theory was also found in two papers. For instance, Zhang et al. [82] was based, specifically, on the kinesthetic learning style theory. Learning styles are the general approaches used by students in learning a new subject [61]. These “overall patterns” that generally direct learning behavior are divided in dimensions, for example, the sensory preference [14]. Sensory preferences can be divided into four main areas: visual, auditory, kinesthetic (movement-oriented)—explored in Zhang et al. [82], and tactile (touch-oriented) [61].

Studio-based learning theory was found only in one study [77]. It is a learning model first developed as part of education and training and later adopted by architectural education in the 1800s [43]. This model has its roots on the notion of the apprentice in the atelier where they worked and learned skills of the master design or artist. Young apprentices did not learn in isolated schools, but were exposed to real adult world and worked on real products in the community.

Other theories were represented by one paper each. Another one was the flow theory that brings the concept of flow which is a state of complete absorption or engagement in an activity that acts as a motivating factor in daily activities such as work, sport, and education [16]. This state encourages a person to persist at an activity due to experience rewards it promises, and it fosters the growth of skills over time [59].

Most of these theories have in common a learner-centered approach, thus focusing more on student’s discovery, construction and interaction process, and the attachment to the context of learning. In this sense, AR, along with other types of technology, can expand the learning horizons. Some theories focus on understanding learning processes to provide a more effective experience for the students considering their personal needs and

abilities. As shown, the trend is to look at AR instructional design from the learners' perspective.

The following question was: "what technologies AR is combined with?". This question inquired if AR applications were combined with other technologies and what kinds of technologies they were combined with. As can be seen in Table 5, 37 papers did not combine AR with other types of technology. The other papers combined it with different types of technology, such as YouTube tutorial, personal blogs, digital sketching, notes and texts provided by the teacher, robotics, mobile pedestrian navigation, virtual reality and digital sketching using hybrid models (DS/HM), and web-based simulation environment. All these technologies appeared one time each. Although it is evident a preference to not combine AR with other types of technology, it is interesting to note that in the classroom environment, AR is another possibility among many others already present in that environment. It is helpful, thus, to understand how these multiple possibilities can work together to scaffold learning.

Question 13 refers to the involvement of teachers in the evaluation process as shown in Fig. 12.

Most studies did not involve the teachers in the studies. Some of the studies were in different contexts, such as library instruction by Wang et al. [74]; thus, in this case, authors mentioned the role of the librarian.

Nevertheless, 13 studies reported the involvement of teachers in different ways and levels. Figure 12 evidences that the teacher may be involved in the design and evaluation process of AR educational tools in different ways. The most common way was the teacher(s), or in some cases, schools directors, working as consultants or curators. Teachers were consulted for different purposes, such

**Table 5** Technologies AR is combined with

Study Reference	AR Combination
[50], [44], [42], [74], [71], [52], [64], [38], [32], [37], [69], [49], [80], [62], [7], [36], [9], [79], [48], [31], [8], [10], [45], [53], [18], [75], [19], [20], [51], [8], [77], [30], [82], [67], [70], [76]	None
[13]	YouTube tutorial
[27]	Personal blogs
[28]	Digital sketching
[17]	Notes and texts provided by the teacher
[33]	Robotics
[39]	Mobile pedestrian navigation
[29]	Virtual Reality and Digital Sketching using Hybrid Models (DS/HM)
[35]	Web-based simulation environment

as problematic contents to teach [82] or to review or modify tests [36, 67, 82].

Seven studies involved the teachers as evaluators of student outputs. As an example, [44] explains that "AR will be used for self-assessment and that the teacher can mark the answers and give the scores on internet web page."

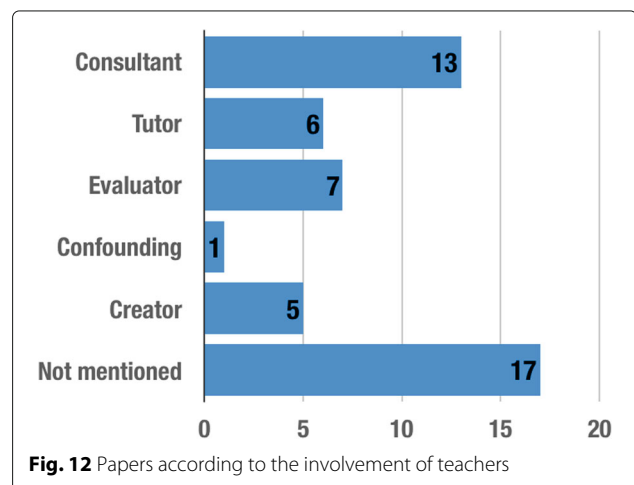
Another role was to act as a tutor (six mentions). That means the teachers had a role of explaining content to students or monitor their work. For instance, [73] mentions that "the procedure of experiment is started with teacher lectures to all students in class."

Also, five papers reported the participation of the teachers as creators of learning experiences. Cubillo et al. [17] reports that the teachers can follow an established procedure to create content using the tool. In da Silva et al. [19] and da Silva et al. [20], teachers were not able to create the applications by themselves since the AR tool evaluated needs an authoring tool, but they were able to design the activities to be worked and programmers created the content accordingly.

Finally, in Frank and Kapila [30], the teacher was considered a confounding, as illustrated in these lines: "teacher's feedback was prevented in the design of the experiment by having student participants tested individually, being directed to perform the activity immediately after the pre-assessment and then immediately to complete the post-assessment."

The results for questions 14 and 15 can be seen in Fig. 13. Q14 refers to the use of multiple metrics. We can see that 24 studies used both quantitative and qualitative metrics and that 21 did not adopt this practice. However, most papers did not use both metrics to evaluate learning gains.

Papers, such as Zhang et al. [82] and Wei et al. [77] used both types of metrics to evaluate learning gains. Zhang et al. [82] investigated the application of location-based AR to astronomical observation instruction. It used both quantitative and qualitative data to investigate aspects



**Fig. 12** Papers according to the involvement of teachers

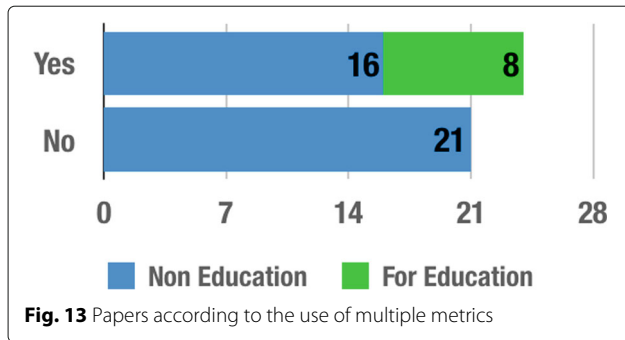


Fig. 13 Papers according to the use of multiple metrics

related to learning. To gather qualitative data, the authors performed an interview with teachers to understand the limitations of traditional teaching methods as a reference for the system’s design proposed. The quantitative data assessed learning effectiveness and motivation.

On the other hand, Wei et al. [77] showed a general technical creative design teaching scheme that includes AR. It used questionnaires to assess creative design learning, motivation, and teaching efficiency. There were also tests on creative design learning motivation, teaching effects, and creativity of the output.

This is an interesting aspect since the educational aspects are very complex and only quantitative metrics are not enough to understand the nuances involved in the process.

**Relation and effect questions**

Question 15 was in this category. This question explores the kind of impact of the tools analyzed in the studies. As shown in Fig. 14, 33 papers reported positive of the results. For instance, Jerry and Aaron [38] proposed a system that promoted a better relation to physics concepts.

Ibáñez et al. [36] revealed that the AR-based application was more effective than the web-based one in promoting student’s knowledge. The four teachers in Wei et al. [77] considered the creative designs produced with AR by students more novel, sophisticated, and with more practical value.

In terms of performance improvement, Yeo et al. [80] reported that the AR image overlay and laser guidance

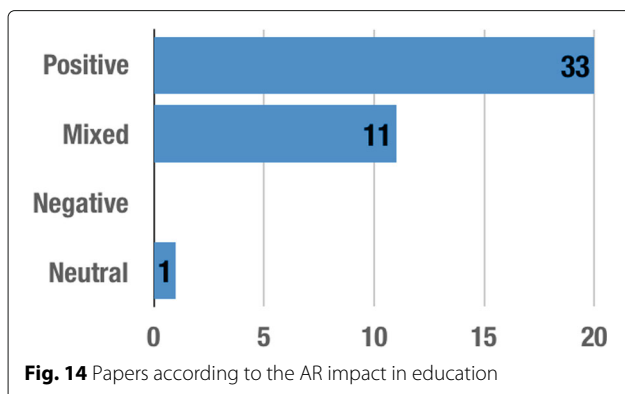


Fig. 14 Papers according to the AR impact in education

improved the training process of needle placement. The participants who trained with overlay guidance performed better even when required to do freehand insertions. Zhang et al. [82] describe that in outdoor teaching environments, altering tool factors significantly enhances performance factors.

Regarding usability aspects, Wei et al. [77] reported that students considered the teaching contents with AR relevant and so had greater satisfaction.

The systems were described as convenient/interesting in some studies. Additionally, students in Tarng et al. [70] considered the virtual scenes and butterflies very realistic, and they would like to use it again in the future.

Reduction in costs were also reported. Student’s attention was also significantly improved due to the introduction of AR technology as reported in Wei et al. [77].

AR also enabled learning formal contents in informal environments as shown in Sommerauer and Müller [67]. This study pointed out that the empirical evidence suggests that AR has the potential to be an effective tool for learning mathematics in a museum. Students also perceived AR as a valuable add-on of the exhibition.

Eleven papers reported mixed results. That means the results could be either positive or negative for one aspect and neutral for others, for example. This situation is illustrated in Martín-Gutiérrez et al. [52]. This paper reported improvement on user’s spatial skills while working on their own; the statistic results show that use of the HMD device does not provide any difference when obtaining spatial ability upgrades with respect to the PC monitor. Authors argue that this result may be caused by the fact that HMD use is not the most suitable as users stated that the glass and camera set were not comfortable.

In Wang et al. [74], the proposed librarian system was more helpful in promoting the learning performance of learners with the field-dependent cognitive style than the conventional librarian instruction, particularly for learning content associated with application and comprehension.

Chen and Tsai [9] revealed that there was no gender difference in learning. This study investigated the AR’s impact depending on student’s personal learning styles (there was an impact) and personal gaming skills (there was no impact). Chen and Tsai [9] revealed a neutral outcome.

Another example is [37], which reported positive regarding intrinsic motivation, but slightly negative (although not significantly different) regarding self-learning. Nevertheless, no paper reported only negative or neutral outcomes.

**Guidelines for educational evaluation**

Through this literature review, authors were able to understand the current status of AR evaluation in

education. In this section, we will discuss some principles that are important to be taken into account in similar situations. These aspects have already been discussed in [18].

Many studies have pointed out the importance of multiple metrics in research design. For instance, Easterbrook et al. [25] point out its usefulness and highlight the importance of employing both quantitative and qualitative metrics as a way of compensating the weakness of each method. Cohen et al. [12] explain that there are many advantages of using multimethod approaches in social research. The authors highlight two of them:

1. While single observation in fields such as physics and chemistry usually yield sufficient and unambiguous information, it provides a limited view of the complexity of human behavior and interactions.
2. Exclusive reliance on one method may bias or distort the researcher's picture of a particular reality he/she is investigating.

Although not all the papers used multiple metrics to evaluate educational aspects, we observed that many papers did use them in their studies.

Another important issue is technology integration into the classrooms. In order to effectively evaluate new educational technology, it is important to effectively integrate them in the schools. Dexter [21] points out two premises for effective integration and implementation of technology for K–12 classrooms, that are:

1. The teacher must act as an instructional designer, planning the use of technology to support learning.
2. Schools must support teachers in this role.

It is important for researchers and developers to have an understanding on how teachers will integrate new technologies into their lessons since this will shape student's learning opportunities. Fitzpatrick [26] stresses the need to involve teachers in the process of adopting new technology, so the activities are integrated to their lesson plan and meaningful to the students. For instance, activity theory [47] shows that activities are culturally mediated and inserted into a given context that includes the mediation of artifacts, of the community, and of its rules and its division of labor. In the process of transforming the activity of teaching into learning, there is a whole complex of mediations involving the curriculum, the educational rules, teacher's training, and artifacts to name a few. This complex scenario needs to be taken into account in order for researchers to understand the changes caused by the introduction of a new artifact and the changes needed to expand and adjust the system.

Hence, taking this information into account, it is possible to infer that teachers need to have a very active approach when it comes to use and evaluation of technology in education. However, the data showed that only

five papers considered the teacher as a creator in their evaluation process.

Crompton [15] explains that the evaluation of a piece of technology in isolation will tend to focus on various aspects of the technology itself, such as screen design and text layout. On the other hand, the evaluation of a courseware within the course itself will allow for examination of other factors that will lead to successful integration of the product within the course. Some of these aspects are:

- Educational setting
- Aims and objectives of the course
- Teaching approach
- Learning strategies
- Assessment methods
- Implementation strategy

Formative evaluations as stated by Scriven are typically conducted during the development or improvement of a program, person, or product, and it is conducted with the intent to improve [66]. On the other hand, summative evaluation is typically quantitative, using numeric scores or letter grades to assess learner achievement. Thus, a comprehensive evaluation involving both types of assessment is advisable in order to have a better overview of the process and its outcome.

### Final remarks

Through this research, we identified AR's potential to be applied in learning contexts. Developments in AR technology have enabled researchers to develop and evaluate more tools in the field of education. Hence, it was evident a growing interest in evaluating its impact in the learning process.

Results have shown that most studies combined different methodologies to evaluate their tools; however, only few papers combined them to evaluate educational gains.

Most of these papers used multiple metrics but to evaluate different aspects rather than just learning, such as usability and efficiency. Merriam [56] explains that all research designs can be discussed in terms of their relative strengths and limitations. She claims that their merits are related to select the most appropriate ones to address the research problem. Cohen et al. [12] argue that there are many advantages of using multimethod approach in social research. They highlight that (a) while single observation in fields such as physics usually yield sufficient and unambiguous information, it provides a limited view of the complexity of human behavior and interactions, and (b) exclusive reliance on one method may bias or distort the researcher's picture of a particular reality.

It was also evident that most studies did not involve the teacher as an instructional designer. However, teachers were involved in many studies in a wide range of ways from consultant to creator. Fitzpatrick [26] highlights the need to involve teachers in the process of adopting new



technological tools, so activities are integrated into their lesson plans and, thus, meaningful to the students.

Although AR has been shown to be helpful for teachers, it can also be inferred that its use, in some situations, may decrease the role of the teacher as the only source of knowledge since it may enable learners to be aided by other peers, trainers, or even their parents depending on the situation.

In this review, we noticed that there are solutions being developed to different age groups and knowledge domains. However, it was noticed a lack of evaluation of AR systems aimed at very young learners.

Regarding the types of questions asked, most papers presented more than one question. These questions were mainly relationship and descriptive-comparative ones. Those papers intended to describe the effect of a given AR technology comparing it with different resources as well as its relationship with different aspects, such as academic achievement or motivation, which indicates that the field is still maturing when it comes to evaluating AR educational impacts.

The papers were also classified according to the tracking, display, and interaction techniques used. It was noticeable that this choice of technology varied deeply depending on the learning objectives of the tool. However, this choice had an impact in the possibilities and limitations of use of the applications.

We also investigated if the papers based their work in any educational theory. Most papers mentioned educational theories. However, 19 studies did not mention any theory. It is important to highlight that educational theories may help to unravel contributions of AR tools as well as its limitations. In addition, it may help to understand how AR unique features may impact in the learning setting. The theories mentioned varied considerably, but something that most of them had in common is a learner-centered approach, thus putting the focus on student's discovery, construction, and interaction processes and the attachment to the learning context.

It is noticeable that AR can expand the learning horizons. Some of the theories focus on understanding learning processes to provide a more effective experience for students considering their personal needs and abilities. We observed the need to look at AR instructional design from the perspective and limitations of the learners themselves.

The latter question investigated the kinds of impact of the results of the studies. Most of them presented positive outcomes. AR has been proved to be a helpful tool concerning many aspects of learning. In this sense, studies presented positive outcomes regarding a wide range of aspects, such as learning, academic performance, and motivation, among others.

Neutral outcomes were also reported as in some studies; the proposed AR system generated equivalent learning performance when compared to a traditional one. However, as already discussed, in many cases, results were neutral for one aspect and positive for others.

The analysis evidenced that AR can help to promote independence and interest among students, which can lead to more student-centered approaches, in which students are the center of their own learning and may apply it in more practical ways. The use of AR also enabled students to experience more concrete situated learning experiences, and together with mobile technologies, it may help to extend learning to different environments in a contextualized way, such as museums and student's campi.

To sum up, during this review, it was noticed that AR has unique affordances that can impact the learning experience. As technology matures, researchers are increasingly concerned with how to incorporate real classroom/learning issues into their investigation.

Thus, authors discussed some guidelines for AR educational evaluation based on the lessons learned. First, based on the literature review, we advocate for the use of multiple metrics both quantitative and qualitative in order to have a better overview of the technology inserted in the teaching context as well as its effects.

Second, although it is not always possible to have a longitudinal evaluation, it is recommended to have a comprehension of more than punctual assessments but rather understand its effect in student's development in a longer term. Finally, as it is widely recognized that teachers play a major role in technology adoption in the schools, we advocate for the involvement of teachers in the evaluation in more active ways as possible. Moreover, it is important to have tools that are flexible enough in order to facilitate teachers' and students' input of content.

As for limitations, due to the limited number of databases used, authors are aware that results may not fully represent the research development in the field.

### Implications of the research

As implications of this research, it was noticed the need for more authoring tools that would enable users to create their own materials independently. Moreover, it is evident the need for more research regarding the evaluation of AR, especially, long-term ones since they could provide a better overview of the process of using this technology into the learning environment.

### Endnotes

<sup>1</sup> <https://www.instructionaldesigncentral.com/whatisinstructionaldesign>

<sup>2</sup> The database with the selected papers is available at: <https://papercatalog.000webhostapp.com>







- C6: Data collection methods clearly described and systematic?  
 C7: Data analysis clearly described and systematic?  
 C8: Use of verification procedure(s) to establish credibility?  
 C9: Conclusions supported by the results?  
 C10: Reflexivity of the account?

#### Acknowledgements

The authors would like to thank Rafael Roberto for his valuable contributions.

#### Funding

The authors would like to thank Fundação de Amparo a Ciência e Tecnologia de Pernambuco (FACEPE) (processes IBPG-0605-1.03/15) for partially funding this research.

#### Availability of data and materials

The datasets supporting the conclusions of this article are available at <https://papercatalog.000webhostapp.com/>.

#### Authors' contributions

MMOS contributed to the conceiving and designing of the research protocol as well as acquiring, analyzing, and interpreting the data and drafting the manuscript. JMXNT contributed to the conceiving and designing the research protocol, analyzing and interpreting the data, and drafting and revising the manuscript. VT and PSC contributed to the conceiving and designing of the research protocol, revising the manuscript, and coordinating the research. All authors read and approved the final manuscript.

#### Authors' information

MMOS is a Ph.D. candidate in Computer Science at the Federal University of Pernambuco and a researcher at Voxar Labs. Her research interests include interactive media, evaluation of educational tools, augmented reality for education and teaching and learning of mother and foreign languages. JMXNT is an assistant professor at Electronics and Systems Department of Federal University of Pernambuco and senior scientist at Voxar Labs. His research interests include 3D tracking, augmented reality, computer vision, computer graphics and embedded systems. VT is an associate professor at the Federal University of Pernambuco and head of the Voxar Labs research group. Her research interests include augmented reality, visualization, tracking and interaction. PSC is an associate professor at the Federal University of Pernambuco and head of the GENTE - New Technologies and Education research group. Her research interests include MOOCS and teachers' long life formations, Learning Spaces and Learning Scenarios.

#### Competing interests

The authors declare that they have no competing interests.

#### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

#### Author details

<sup>1</sup>Voxar Labs, Centro de Informática, Universidade Federal de Pernambuco, Recife, Brazil. <sup>2</sup>Departamento de Eletrônica e Sistemas, Universidade Federal de Pernambuco, Recife, Brazil. <sup>3</sup>EDUMATEC, Centro de Educação, Universidade Federal de Pernambuco, Recife, Brazil.

Received: 11 April 2018 Accepted: 10 January 2019

Published online: 04 February 2019

#### References

- Altman D (1990) Practical statistics for medical research. Chapman & Hall/CRC Texts in Statistical Science. Taylor & Francis, London. ISBN 9780412276309. <https://books.google.com.br/books?id=v-walRnRxWQC>
- Amor B (2014) Experience art through relative realities with Samsung digital gallery at Yuchengco Museum. <https://goo.gl/KLygPJ>. Accessed 2 Oct 2018
- Azuma RT (1997) A survey of augmented reality. Presence: Teleoper. Virtual Environ 6(4):355–385. ISSN 1054-7460. <https://doi.org/10.1162/pres.1997.6.4.355>.
- Balog A, Pribeanu C (2010) The role of perceived enjoyment in the students' acceptance of an augmented reality teaching platform: a structural equation modelling approach. Stud Inform Control 19(3):319–330
- Billinghurst M, Duenser A (2012) Augmented reality in the classroom. Computer 45(7):56–63. ISSN 0018-9162. <https://doi.org/10.1109/MC.2012.111>
- Billinghurst M, Clark A, Lee G, et al (2015) A survey of augmented reality. Found Trends® Hum-Comput Interact 8(2-3):73–272
- Bosque LD, Martinez R, Torres JL (2015) Decreasing failure in programming subject with augmented reality tool. Procedia Comput Sci 75:221–225. ISSN 1877-0509. <https://doi.org/10.1016/j.procs.2015.12.241>. <http://www.sciencedirect.com/science/article/pii/S1877050915037023>
- Chang K-E, Chang C-T, Hou H-T, Sung Y-T, Chao H-L, Lee C-M (2014) Development and behavioral pattern analysis of a mobile guide system with augmented reality for painting appreciation instruction in an art museum. Comput Educ 71:185–197. ISSN 0360-1315. <https://doi.org/10.1016/j.compedu.2013.09.022>. <http://www.sciencedirect.com/science/article/pii/S0360131513002868>
- Chen C-M, Tsai Y-N (2012) Interactive augmented reality system for enhancing library instruction in elementary schools. Comput Educ 59(2):638–652. ISSN 0360-1315. <https://doi.org/10.1016/j.compedu.2012.03.001>. <http://www.sciencedirect.com/science/article/pii/S0360131512000589>
- Cheng K-H, Tsai C-C (2014) Children and parents' reading of an augmented reality picture book: analyses of behavioral patterns and cognitive attainment. Comput Educ 72:302–312. ISSN 0360-1315. <https://doi.org/10.1016/j.compedu.2013.12.003>. <http://www.sciencedirect.com/science/article/pii/S0360131513003278>
- Cohen J (1960) A coefficient of agreement for nominal scales. Educ Psychol Meas 20(1):37–46
- Cohen L, Manion L, Morrison K (2003) Research methods in education. Taylor & Francis ebook collection. Taylor & Francis. ISBN 9780203224342. <https://books.google.co.uk/books?id=5twk1pHwyl8C>. Accessed 2 Oct 2018
- Contero M, Gomis JM, Naya F, Albert F, Martin-Gutierrez J (2012) Development of an augmented reality based remedial course to improve the spatial ability of engineering students. In: 2012 Frontiers in Education Conference Proceedings. pp 1–5. <https://doi.org/10.1109/FIE.2012.6462312>
- Cornett CE (1983) What you should know about teaching and learning styles. Fastback Series. In: Phi Delta Kappa Educational Foundation. ISBN 9780873671910. <https://books.google.at/books?id=SlckAQAAMAAJ>. Accessed 2 Oct 2018
- Crompton P (1996) Evaluation: a practical guide to methods. In: Learn Technol Dissemination Initiative. p 66
- Csikszentmihalyi M (1990) Flow: the psychology of optimal experience, 6 edn. Harper & Row, New York
- Cubillo J, Martín S, Castro M, Diaz G, Colmenar A, Boticki I (2014) A learning environment for augmented reality mobile learning. In: 2014 IEEE Frontiers in Education Conference (FIE) Proceedings. pp 1–8. <https://doi.org/10.1109/FIE.2014.7044039>
- da Silva MMO, Roberto RA, Teichrieb V, Cavalcante PS (2016) Towards the development of guidelines for educational evaluation of augmented reality tools. In: 2016 IEEE Virtual Reality Workshop on K-12 Embodied Learning through Virtual Augmented Reality (KELVAR). pp 17–21. <https://doi.org/10.1109/KELVAR.2016.7563677>
- da Silva M, Roberto R, Teichrieb V (2013) Evaluating an educational system based on projective augmented reality. In: Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE), vol 24. Sociedade Brasileira de Computação, Porto Alegre. p 214
- da Silva MMO, Roberto R, Teichrieb V (2015) Evaluation of augmented reality technology in the English language field. In: Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE), vol 26. Sociedade Brasileira de Computação, Porto Alegre. p 577
- Dexter S (2002) eTIPS-Educational technology integration and implementation principles. Designing Instr Technol-Enhanc Learn 24:56–70
- Dünser A, Billinghurst M (2011) Evaluating augmented reality systems. Springer New York, New York. pp 289–307. ISBN 978-1-4614-0064-6. [https://doi.org/10.1007/978-1-4614-0064-6\\_13](https://doi.org/10.1007/978-1-4614-0064-6_13)

23. Dünser A, Grasset R, Seichter H, Billingham M (2007) Applying HCI principles to AR systems design. Technical report. HIT Lab NZ
24. Dünser A, Grasset R, Billingham M (2008) A survey of evaluation techniques used in augmented reality studies. In: ACM SIGGRAPH ASIA 2008 courses, SIGGRAPH Asia '08. ACM, New York. pp 1–27. <https://doi.org/10.1145/1508044.1508049>. <http://portal.acm.org/citation.cfm?id=1508044.1508049>
25. Easterbrook S, Singer J, Storey M-A, Damian D (2008) Guide to advanced empirical software engineering, chapter Selecting Empirical Methods for Software Engineering Research. Springer London, London. ISBN 978-1-84800-044-5. [https://doi.org/10.1007/978-1-84800-044-5\\_11](https://doi.org/10.1007/978-1-84800-044-5_11). [http://dx.doi.org/10.1007/978-1-84800-044-5\\_11](http://dx.doi.org/10.1007/978-1-84800-044-5_11)
26. Fitzpatrick A (2004) Analytical survey information and communication technologies in the teaching and learning of foreign languages: state-of-the-art, needs and perspectives. unesco institute for information technologies in education. Technical report. UNESCO, Moscow, Russia
27. Fonseca D, Villagrasa S, Valls F, Redondo E, Climent A, Vicent L (2014) Engineering teaching methods using hybrid technologies based on the motivation and assessment of student's profiles. In: 2014 IEEE Frontiers in Education Conference (FIE) Proceedings. pp 1–8. <https://doi.org/10.1109/FIE.2014.7044209>
28. Fonseca D, Villagrasa S, Valls F, Redondo E, Climent A, Vicent L (2014) Motivation assessment in engineering students using hybrid technologies for 3d visualization. In: 2014 International Symposium on Computers in Education (SIIE). pp 111–116. <https://doi.org/10.1109/SIIE.2014.7017714>
29. Fonseca D, Valls F, Redondo E, Villagrasa S (2016) Informal interactions in 3D education: Citizenship participation and assessment of virtual urban proposals. *Comput Hum Behav* 55:504–518. ISSN 0747-5632. <https://doi.org/10.1016/j.chb.2015.05.032>. <http://www.sciencedirect.com/science/article/pii/S0747563215004033>
30. Frank JA, Kapila V (2017) Mixed-reality learning environments: integrating mobile interfaces with laboratory test-beds. *Comput Educ* 110:88–104. ISSN 0360-1315. <https://doi.org/10.1016/j.compedu.2017.02.009>. <http://www.sciencedirect.com/science/article/pii/S0360131517300325>
31. Hou L, Wang X (2013) A study on the benefits of augmented reality in retaining working memory in assembly tasks: a focus on differences in gender. *Autom Constr* 32:38–45. ISSN 0926-5805. <https://doi.org/10.1016/j.autcon.2012.12.007>. <http://www.sciencedirect.com/science/article/pii/S0926580512002464>
32. Hsiao KF, Rashvand HF (2011) Body language and augmented reality learning environment. In: 2011 Fifth FTRA International Conference on Multimedia and Ubiquitous Engineering. pp 246–250. <https://doi.org/10.1109/MUE.2011.51>
33. Hulin T, Schmirgel V, Yechiam E, Zimmermann UE, Preusche C, Pöhler G (2010) Evaluating exemplary training accelerators for programming-by-demonstration. In: 19th International Symposium in Robot and Human Interactive Communication. pp 440–445. <https://doi.org/10.1109/ROMAN.2010.5598611>
34. Hutchings W (2007) Enquiry-based learning: definitions and rationale. Technical report, Centre for Excellence in Enquiry-Based. [http://www.ceebl.manchester.ac.uk/resources/papers/hutchings2007\\_definingebl.pdf](http://www.ceebl.manchester.ac.uk/resources/papers/hutchings2007_definingebl.pdf)
35. Ibáñez MB, Di-Serio Á, Villarán-Molina D, Delgado-Kloos C (2015) Augmented reality-based simulators as discovery learning tools: an empirical study. *IEEE Trans Educ* 58(3):208–213. ISSN 0018-9359. <https://doi.org/10.1109/TE.2014.2379712>
36. Ibáñez MB, Di Serio Á, Villarán D, Kloos CD (2014) Experimenting with electromagnetism using augmented reality: impact on flow student experience and educational effectiveness. *Comput Educ* 71:1–13. ISSN 0360-1315. <https://doi.org/10.1016/j.compedu.2013.09.004>. <http://www.sciencedirect.com/science/article/pii/S0360131513002571>
37. Iwata T, Yamabe T, Nakajima T (2011) Augmented reality go: extending traditional game play with interactive self-learning support. In: 2011 IEEE 17th International Conference on Embedded and Real-Time Computing Systems and Applications, vol 1. pp 105–114. <https://doi.org/10.1109/RTCSA.2011.43>
38. Jerry TFL, Aaron CCE (2010) The impact of augmented reality software with inquiry-based learning on students' learning of kinematics graph. In: 2010 2nd International Conference on Education Technology and Computer, vol 2. pp V2–1–V2–5. <https://doi.org/10.1109/ICETC.2010.5529447>
39. Joo-Nagata J, Abad FM, Giner JG-B, García-Peñalvo FJ (2017) Augmented reality and pedestrian navigation through its implementation in m-learning and e-learning: evaluation of an educational program in Chile. *Comput Educ* 111:1–17. ISSN 0360-1315. <https://doi.org/10.1016/j.compedu.2017.04.003>. <http://www.sciencedirect.com/science/article/pii/S0360131517300751>
40. Kmet LM, Lee RC, AHFMR, Cook LS (2004) Standard quality assessment criteria for evaluating primary research papers from a variety of fields. In: HTA initiative. Alberta Heritage Foundation for Medical Research. ISBN 9781896956770. <https://books.google.at/books?id=UYE8HQAAACAJ>. Accessed 2 Oct 2018
41. Kostaras NN, Xenos MN, et al (2009) Assessing the usability of augmented reality systems. In: Proceedings of the 13th Panhellenic Conference on Informatics, Alberta Heritage Foundation for Medical Research (AHFMR). AHFMR-HTA Initiative. pp 197–201
42. Kraut B, Jeknić J (2015) Improving education experience with augmented reality (AR). In: 2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO). pp 755–760. <https://doi.org/10.1109/MIPRO.2015.7160372>
43. Lackney JA (1999) A history of the studio-based learning model. [http://edi.msstate.edu/work/pdf/history\\_studio\\_based\\_learning.pdf](http://edi.msstate.edu/work/pdf/history_studio_based_learning.pdf). Accessed 2 Oct 2018
44. Lai ASY, Wong CYK, Lo OCH (2015) Applying augmented reality technology to book publication business. In: 2015 IEEE 12th International Conference on e-Business Engineering. pp 281–286. <https://doi.org/10.1109/ICEBE.2015.55>
45. Leblanc F, Champagne BJ, Augestad KM, Neary PC, Senagore AJ, Ellis CN, Delaney CP (2010) A comparison of human cadaver and augmented reality simulator models for straight laparoscopic colorectal skills acquisition training. *J Am Coll Surg* 211(2):250–255. ISSN 1072-7515. <https://doi.org/http://doi.org/10.1016/j.jamcollsurg.2010.04.002>. <http://www.sciencedirect.com/science/article/pii/S1072751510002735>
46. Lee K (2012) Augmented reality in education and training. *TechTrends* 56(2):13–21. ISSN 1559-7075. <https://doi.org/10.1007/s11528-012-0559-3>. <http://dx.doi.org/10.1007/s11528-012-0559-3>
47. Leont'ev AN (1978) Activity, consciousness, and personality. Prentice-Hall, Englewood Cliffs. ISBN 9780130035332. <http://chc.ucsd.edu/MCA/Paper/leontev/>
48. Lin T-J, Duh HB-L, Li N, Wang H-Y, Tsai C-C (2013) An investigation of learners' collaborative knowledge construction performances and behavior patterns in an augmented reality simulation system. *Comput Educ* 68:314–321. ISSN 0360-1315. <https://doi.org/10.1016/j.compedu.2013.05.011>. <http://www.sciencedirect.com/science/article/pii/S0360131513001322>
49. Mahmoudi MT, Badie K, Valipour M (2015) Assessing the role of AR-based content in improving learning performance considering Felder-Silverman learning style. In: 2015 International Conference on Interactive Collaborative Learning (ICL). pp 838–843. <https://doi.org/10.1109/ICL.2015.7318137>
50. Martín-Gutiérrez J (2011) Proposal of methodology for learning of standard mechanical elements using augmented reality. In: 2011 Frontiers in Education Conference (FIE). pp 1–6. <https://doi.org/10.1109/FIE.2011.6142708>
51. Martín-Gutiérrez J, Saorin JL, Contero M, Alcaniz M (2010) Ar\_Dehaes: an educational toolkit based on augmented reality technology for learning engineering graphics. In: 2010 10th IEEE International Conference on Advanced Learning Technologies. pp 133–137. <https://doi.org/10.1109/ICALT.2010.45>
52. Martín-Gutiérrez J, Navarro RE, González MA (2011) Mixed reality for development of spatial skills of first-year engineering students. In: 2011 Frontiers in Education Conference (FIE). pp T2D–1–T2D–6. <https://doi.org/10.1109/FIE.2011.6142707>
53. Martínez AA, Benito JRL, González EA, Ajuria EB (2017) An experience of the application of augmented reality to learn english in infant education. In: 2017 International Symposium on Computers in Education (SIIE). pp 1–6. <https://doi.org/10.1109/SIIE.2017.8259645>
54. Mayer RE (2009) Multimedia learning, 2 edn. Cambridge University Press, Cambridge
55. McGreal R (2009) Mobile devices and the future of free education. [http://www.ou.nl/Docs/Campagnes/ICDE2009/Papers/Final\\_paper\\_252mcgreal.pdf](http://www.ou.nl/Docs/Campagnes/ICDE2009/Papers/Final_paper_252mcgreal.pdf). Accessed 07 Oct 2018

56. Merriam SB (2009) Qualitative research: a guide to design and implementation. Jossey-Bass higher and adult education series. John Wiley & Sons. ISBN 9780470283547. <https://books.google.at/books?id=tvFICrgcuSIC>. Accessed 2 Oct 2018
57. Moher D, Liberati A, Tetzlaff J, Altman DG (2010) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Int J Surg* 8(5):336–341. ISSN 1743-9191. <https://doi.org/10.1016/j.ijso.2010.02.007>. <http://www.sciencedirect.com/science/article/pii/S1743919110000403>
58. Moura AMC (2010) Apropriação do Telemóvel como Ferramenta de Mediação em Mobile Learning: Estudos de Caso em Contexto Educativo. PhD thesis. Universidade do Minho. <http://repositorium.sdum.uminho.pt/handle/1822/13183>. Accessed 2 Oct 2018
59. Nakamura J, Csikszentmihalyi M (2014) The concept of flow. Springer Netherlands, Dordrecht. ISBN 978-94-017-9088-8. [https://doi.org/10.1007/978-94-017-9088-8\\_16](https://doi.org/10.1007/978-94-017-9088-8_16).
60. Ontario Ministry of Education (2013) Inquiry-based learning. Technical report, Ontario Ministry of Education. [http://www.edu.gov.on.ca/eng/literacynumeracy/inspire/research/CBS\\_InquiryBased.pdf](http://www.edu.gov.on.ca/eng/literacynumeracy/inspire/research/CBS_InquiryBased.pdf)
61. Oxford RL (2001) Learning styles and strategies. In: Celce-Murcia M (ed). *Teaching English as a Second or Foreign Language*. Heinle Cengage Learning, Boston. pp 359–366
62. Ramírez H, Mendoza E, Mendoza M, González E (2015) Application of augmented reality in statistical process control, to increment the productivity in manufacture. *Procedia Comput Sci* 75:213–220. ISSN 1877-0509. <https://doi.org/10.1016/j.procs.2015.12.240>. <http://www.sciencedirect.com/science/article/pii/S1877050915037011>
63. Roberto R, Lima JP, Teichrieb V (2016) ISSN 0097-8493. <https://doi.org/http://dx.doi.org/10.1016/j.cag.2016.02.002>. <http://www.sciencedirect.com/science/article/pii/S0097849316300115>. *Comput Graph* 56:20–30
64. Salazar M, Gaviria J, Laorden C, Bringas PG (2013) Enhancing cybersecurity learning through an augmented reality-based serious game. In: 2013 IEEE Global Engineering Education Conference (EDUCON). pp 602–607. <https://doi.org/10.1109/EduCon.2013.6530167>
65. Santos MEC, Chen A, Taketomi T, Yamamoto G, Miyazaki J, Kato H (2014) Augmented reality learning experiences: survey of prototype design and evaluation. *IEEE Trans Learn Technol* 7(1):38–56. ISSN 1939-1382. <https://doi.org/10.1109/TLT.2013.37>
66. Scriven M (1991) Beyond formative and summative evaluation. In: *Evaluation and Education: A Quarter Century*
67. Sommerauer P, Müller O (2014) Augmented reality in informal learning environments: a field experiment in a mathematics exhibition. *Comput Educ* 79:59–68. ISSN 0360-1315. <https://doi.org/10.1016/j.compedu.2014.07.013>. <http://www.sciencedirect.com/science/article/pii/S036013151400164X>
68. Swan JE, Gabbard JL (2009) Survey of user-based experimentation in augmented reality. In: 1st International Conference on Virtual Reality, HCI International. pp 1–9. <http://people.cs.vt.edu/jgabbard/publications/hcivr05.pdf>. Accessed 2 Oct 2018
69. Tarrg W, Ou KL (2012) A study of campus butterfly ecology learning system based on augmented reality and mobile learning. In: 2012 IEEE Seventh International Conference on Wireless, Mobile and Ubiquitous Technology in Education. pp 62–66. <https://doi.org/10.1109/WMUTE.2012.17>
70. Tarrg W, Yu CS, Liou FL, Liou HH (2013) Development of a virtual butterfly ecological system based on augmented reality and mobile learning technologies. In: 2013 9th International Wireless Communications and Mobile Computing Conference (IWCMC). pp 674–679. <https://doi.org/10.1109/IWCMC.2013.6583638>
71. Tobar-Muñoz H, Fabregat R, Baldiris S (2014) Using a videogame with augmented reality for an inclusive logical skills learning session. In: 2014 International Symposium on Computers in Education (SIIE). pp 189–194. <https://doi.org/10.1109/SIIE.2014.7017728>
72. Tori R, Kirner C, Siscoutto RA (2006) Fundamentos e tecnologia de realidade virtual e aumentada. Editora SBC. ISBN 9788576690689. <https://books.google.com.br/books?id=h1H7PgAACAAJ>. Accessed 2 Oct 2018
73. Tsai CH, Huang JY (2014) A mobile augmented reality based scaffolding platform for outdoor fieldtrip learning. In: 2014 IIAI 3rd International Conference on Advanced Applied Informatics. pp 307–312. <https://doi.org/10.1109/IIAI-AAI.2014.70>
74. Wang YS, Chen CM, Hong CM, Tsai YN (2013) Interactive augmented reality game for enhancing library instruction in elementary schools. In: 2013 IEEE 37th Annual Computer Software and Applications Conference Workshops. pp 391–396. <https://doi.org/10.1109/COMPASACW.2013.128>
75. Wang Y-H (2017) Exploring the effectiveness of integrating augmented reality-based materials to support writing activities. *Comput Educ* 113:162–176. ISSN 0360-1315. <https://doi.org/10.1016/j.compedu.2017.04.013>. <http://www.sciencedirect.com/science/article/pii/S0360131517300969>
76. Wei L, Najdovski Z, Abdelrahman W, Nahavandi S, Weisinger H (2012) Augmented optometry training simulator with multi-point haptics. In: 2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC). pp 2991–2997. <https://doi.org/10.1109/ICSMC.2012.6378250>
77. Wei X, Wang D, Liu Y, Wang Y (2015) Teaching based on augmented reality for a technical creative design course. *Comput Educ* 81:221–234. ISSN 0360-1315. <https://doi.org/10.1016/j.compedu.2014.10.017>. <http://www.sciencedirect.com/science/article/pii/S0360131514002371>
78. Wei H-K, Lee SW-Y, Chang H-Y, Liang J-C (2013) Current status, opportunities and challenges of augmented reality in education. *Comput Educ* 62(0):41–49. ISSN 0360-1315. <https://doi.org/10.1016/j.compedu.2012.10.024>. <http://www.sciencedirect.com/science/article/pii/S0360131512002527>
79. Yen J-C, Tsai C-H, Wu M (2013) Augmented reality in the higher education: students' science concept learning and academic achievement in astronomy. *Procedia-Soc Behav Sci* 103:165–173. ISSN 1877-0428. <https://doi.org/10.1016/j.sbspro.2013.10.322>. <http://www.sciencedirect.com/science/article/pii/S1877042813037671>
80. Yeo CT, Ungi T, U-Thainual P, Lasso A, McGraw RC, Fichtinger G (2011) The effect of augmented reality training on percutaneous needle placement in spinal facet joint injections. *IEEE Trans Biomed Eng* 58(7):2031–2037. ISSN 0018-9294. <https://doi.org/10.1109/TBME.2011.2132131>
81. Yin RK (2009) Case study research: design and methods. Applied Social Research Methods. SAGE Publications. ISBN 9781412960991. <https://books.google.com.br/books?id=FzawlAdiHkC>
82. Zhang J, Sung Y-T, Hou H-T, Chang K-E (2014) The development and evaluation of an augmented reality-based armillary sphere for astronomical observation instruction. *Comput Educ* 73:178–188. ISSN 0360-1315. <https://doi.org/10.1016/j.compedu.2014.01.003>. <http://www.sciencedirect.com/science/article/pii/S036013151400013X>. Accessed 2 Oct 2018
83. Zhou F, Duh HB-L, Billingham M (2008) Trends in augmented reality tracking, interaction and display: a review of ten years of ISMAR. In: *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*. IEEE Computer Society, Cambridge. pp 193–202

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](http://springeropen.com)