



Review

A systematic review of foreign language learning with immersive technologies (2001-2020)

Rebecca M. Hein^{1,2,*}, Carolin Wienrich² and Marc E. Latoschik¹

¹ Human-Computer Interaction, Julius-Maximilians-Universität Würzburg, Am Hubland, D-97074 Würzburg

² Human-Technique Systems, Julius-Maximilians-Universität Würzburg, Oswald-Külpe-Weg 82, D-97074 Würzburg

* **Correspondence:** Email: rebecca.hein@uni-wuerzburg.de.

Abstract: This study provides a systematic literature review of research (2001–2020) in the field of teaching and learning a foreign language and intercultural learning using immersive technologies. Based on 2507 sources, 54 articles were selected according to a predefined selection criteria. The review is aimed at providing information about which immersive interventions are being used for foreign language learning and teaching and where potential research gaps exist. The papers were analyzed and coded according to the following categories: (1) investigation form and education level, (2) degree of immersion, and technology used, (3) predictors, and (4) criterions. The review identified key research findings relating the use of immersive technologies for learning and teaching a foreign language and intercultural learning at cognitive, affective, and conative levels. The findings revealed research gaps in the area of teachers as a target group, and virtual reality (VR) as a fully immersive intervention form. Furthermore, the studies reviewed rarely examined behavior, and implicit measurements related to inter- and trans-cultural learning and teaching. Inter- and transcultural learning and teaching especially is an underrepresented investigation subject. Finally, concrete suggestions for future research are given. The systematic review contributes to the challenge of interdisciplinary cooperation between pedagogy, foreign language didactics, and Human-Computer Interaction to achieve innovative teaching-learning formats and a successful digital transformation.

Keywords: foreign language learning and teaching; intercultural learning and teaching; immersive technologies; education; human-computer interaction; systematic literature review

1. Introduction

“Digitalization [...] is an issue that concerns many educational stakeholders” [8, p. 1] Digital devices should be commonplace in all schools. However, many scientists, as well as the pandemic, have revealed many challenges of the appropriate digitalization in educational systems [89, 90]. Firstly, the digital equipment of schools is one success factor. Second, the acceptance among teachers and digital competences to handle the devices are crucial success factors. Third, interdisciplinary incorporation of digital and pedagogical knowledge resulting in innovative teaching-learning formats are essential for a successful digital transformation and effective teaching and learning effects [97]. Particularly, immersive technologies provide much potential to train competencies and increase acceptance on the one hand [72, 73]. Concurrently, immersive technologies are geared for the development of innovative teaching-learning formats and effective teaching and learning effects [96]. While Garzón, and Acevedo analyzed 64 quantitative research studies from 2010 to 2018 to investigate the impact of AR on students’ learning outcomes and illustrated that AR had a meaningful effect on student learning gains, there is less research systemizing learning with VR than AR [34]. Hamilton and colleagues noted in their literature review that affective behavioral changes in non-pedagogical VR applications are extensively studied. However, they were underrepresented in education which is an important area for future research [39]. VR can make an essential contribution to education by allowing students to directly experience environments or situations that are difficult to recreate with traditional teaching methods [79, 97]. These potentials might be highly relevant for foreign languages, especially in the current Covid-19 pandemic and climate crises, where immersion programs are less feasible. In sum, the field lacks systematic reviews revealing successful gestalt features of the whole immersive technology spectrum in foreign language learning and research gaps that should be addressed by empirical studies. The present systematic review of the literature addresses foreign language learning with immersive technology to identify and analyze research gaps and encourage the interdisciplinary cooperation between pedagogy, foreign language didactics, and Human-Computer Interaction. Five research questions are to be answered: (RQ1) How are virtual, fully immersive learning environments used for foreign language learning, and (RQ2) which characteristics of immersive technology support foreign language learning? (RQ3) Can virtual, fully immersive learning environments increase motivation, and success in learning a foreign language? (RQ4) Can they change participants’ attitudes through intercultural encounters, and (RQ5) how are they used for teacher training?

2. Background

2.1. A short view on virtual reality

Virtual reality (VR) and augmented reality (AR) are subordinate to the term immersive technologies or extended reality (XR) [99]. AR refers to the real-time combination of digital and physical information through different technological devices, and VR creates a complete, artificial virtual environment and thus offers complete virtualization [9, 99]. In the context of behavior change, the BehaveFIT framework described four potentials of VR to support behavior change processes: self-representation (e.g., perspective-taking), context-representation (e.g., showing different environments), representation of others (e.g., showing diverse others), and the representation of objects (e.g., including personal everyday objects) [96]. Those potentials can be transferred to a teaching and learning context

and have already shown significant influence [1, 2, 44, 74, 77].

Experiencing VR is strongly linked with the concepts of immersion and presence (e.g., [83]). Immersion and presence can probably be considered a so-called hygiene factor [95]. To a certain extent, it might be necessary to allow other VR potentials to become effective. However, the term immersion is understood differently in many areas and also within the field of human-computer interaction [85]. The best known and most widely accepted definitions of immersion in the field of Human-Computer Interaction (HCI) are those of Witmer and Singer [98] and those of Slater [84]. We followed the latter, defining immersion as objective system factors that influence the sense of presence. Immersion stands for what the technology delivers in all sensory and tracking modalities and that it can be objectively assessed (for more discussion about different uses of the term immersion, see, e.g., [83, 95]). In contrast, presence can be defined as a human reaction to a system of a certain level of immersion and thus describes a subjective state [85, 86]. Defining immersion by technological factors, immersive applications can be categorized by the degree of immersion. One possibility is a categorization along the virtual-reality-continuum [69]. VR applications create an entirely artificial virtual environment and pertain as most immersive [9] followed by mixed reality (MR) applications. MR refers to a specific subset of immersive technologies that involve merging real and virtual worlds [69]. This includes, for example, augmented reality (AR) applications and applications in which 360° content is shown through a head-mounted display (HMD). Another wider categorization possibility refers to extended reality (XR) that combines all real and virtual environments and refers to human-machine interactions generated by computer technology and wearables [32]. It is used to cover all these approaches or to indicate their combined use. Therefore, multi-user virtual environments (MUVE's), such as Second Life, can be included in such a categorization definition. To enable a comprehensive systematization, the present review included all studies designated as immersive and follows the wider possibility of categorization.

2.2. A short view on language learning and beyond

Language learning is initially shaped by learning vocabulary, grammar, and text analysis. However, in Europe, where constructivist views on language learning and education dominate the academic discourse, language learning is integrated into communicative-based learning. Language acquisition takes place during a communicative exchange [37, 57]. Topics are mostly related to the foreign country conveying knowledge about culture, traditions, living, for example. Further, due to the advancing globalization, not only language competencies, but also intercultural competencies have become an important part of modern foreign language teaching [33]. While intercultural competence is often seen as a linear model [40], Chen, and Starosta assume a multidimensional construct, “consisting of three interrelated aspects: intercultural sensitivity (affective aspect), intercultural awareness (cognitive aspect) and intercultural adroitness (behavioral aspect)” [13, 103]. However, research shows that many adult students do not possess the necessary intercultural skills [103]. A lack of intercultural competencies leads to prejudice, discrimination, and unfriendly expressions directly related to misunderstandings between people from different cultural backgrounds and affiliations [4] – consequences that might be particularly severe in school. A meta-analysis that examined different interventions for intercultural competencies found that immersion programs moderated intercultural competence more strongly compared to educational interventions with students. The study revealed that immersion programs primarily addresses the cognitive and affective levels. Therefore, in future studies, they recommend to include formats addressing the conative level as well [103]. This

recommendation is in line with modern pedagogic principles focusing on competence- and action-orientated as well as situated learning, connectivity, and co-construction [55].

According to the potentials described above, VR enables an active exploration of distant and diverse learning contexts, the highlighting of achieved milestones, the possibility to change roles, and to encounter the unknown in a tangible way [97]. Thus, the potentials of immersive technologies draw on these pedagogic principles.

Especially in times of Covid-19, where immersion programs are much less feasible and in times of climate crisis, social VR applications may offer enormous potential. Exchange programs and the experience they provide learners with are probably irreplaceable. However, for larger school classes or to create social fairness, a virtual exchange via a social VR application may be the ideal solution, or addendum.

3. Research questions

VR has many potentials for language and intercultural learning. However, the field lacks systematic reviews describing VR interventions, manipulations, distinctive immersion factors, and corresponding outcomes leading to the following research questions:

- | | |
|------------|---|
| RQ1 | How are virtual, fully immersive learning environments used for foreign language learning? |
| RQ2 | Which characteristics of immersive technology support foreign language learning? |
| RQ3 | Can virtual, fully immersive learning environments increase motivation and success in learning a foreign language? |
| RQ4 | Can virtual, fully immersive learning environments change participants' attitudes through intercultural encounters? |
| RQ5 | How are virtual, fully immersive learning environments used for teacher training? |

4. Method

For this literature research and review, we followed the PRISMA statement [70] and the search strategy of Cooper et al. [24]. In addition, we followed the Population Intervention Comparison Outcome (PICO) scheme [61]. The selected populations were schools, universities, or other educational institutes. All interventions analyzed included an immersive setting, i.e., AR, MR, VR, or applications described as immersive by the authors. These immersive interventions should be compared to conventional learning and teaching, or between. The desired outcome is an improvement of language skills and/or an increase in intercultural competence, so a change in attitude towards the foreign culture, their people, or language [10].

In the following, we describe all sources of information used during the search process. Furthermore, we described the selection process of studies (identification, screening, eligibility, inclusion in the systematic review). The number of pre-selected studies, screened for eligibility, and included in the review, with reasons for exclusion at each stage, are shown and presented in a flowchart

(Figure 1). Next, the categories according to which the studies will be examined and classified will be presented to identify successful gestalt features, answers to the research questions, and potential research gaps.

4.1. Data collection

The selection process of studies followed four steps: (1) identification, (2) screening, (3) eligibility, and (4) final inclusion (Figure 1).

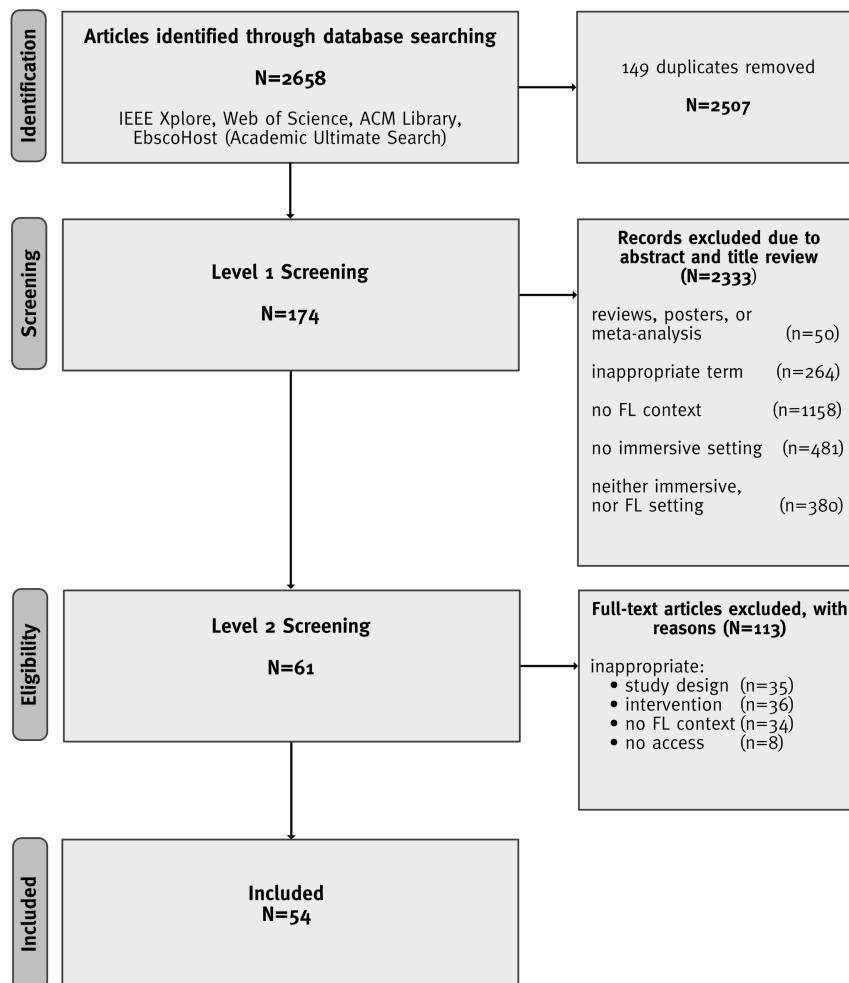


Figure 1. The PRISMA flow diagram for the systematic review detailing the database searches, the number of abstracts screened, the full texts retrieved, and the number of papers included.

STEP 1: identification

For the literature search, the databases IEEE Xplore, EBSCOhost (academic search ultimate), Web of Science, and ACM Library were scanned. The search phrases used resulted in 32 search queries (Table 1). Only peer-reviewed academic journals, conference papers, reports, and reviews that have been published since April 2001, are written in English or German, and contain either qualitative

or quantitative data acquisition or both were included. The article must be an empirical study and explicitly address the learning of a second language. We did not explicitly search for intercultural learning since our focus was on intercultural aspects during language learning. The last search request was submitted on May 31, 2020. The studies' intervention had to include an immersive setting, e.g., AR, MR, VR, or a system described as immersive by the authors. All papers that had unfitting terms or did not fit the journal topic or population were excluded. The search yielded 2507 results after removing duplicates.

Table 1. The table shows the defined key words and their order in the individual search queries.

1st keyword	AND	2nd keyword	AND	3rd keyword	
foreign language		mixed reality	OR	learning	OR
		virtual reality	OR	teaching	OR
		augmented reality	OR	perspective	OR
		immers*	OR	education*	OR
				virtual exchange	OR
				intercult*	OR
				attitud*	OR
				chang*	OR

STEP 2: screening

During the Level 1 screening, the titles and abstracts of the papers were scanned. Studies without qualitative, quantitative, or mixed-methods data collection were excluded. In addition, papers were excluded whose target group was not appropriate (n=18). These included, for instance, studies with the following terms: blind, autism, deaf, disabilit*, down syndrome, etc. People with disabilities have special needs for learning a language, which we do not want to investigate in this context. Given the frequent overlap of these excluded target groups with unrelated terms related to “medical” these excluded studies were included in this category. Further, unrelated topics included papers with medical terms like “rehabilitation”, “medical”, or “physio*”, but also general medical journals (n=97), with industrial terms like “Internet of Things”, or “mechanical” (n=61), or with terms like “climate change”. Moreover, we noticed that immersion was often used in connection with an immersion program or immersion class. That implied it was about an exchange program and not about immersion through a technical medium. These papers were excluded, as well (n=481). In connection with the term “behavior, attitude, or change” and the 2nd keywords of the search query, many papers retrieved were then excluded, because no educational topic was apparent (n=69). In most cases, the papers were not related to foreign language (FL) learning, or teaching (n=1158). There were also papers in which neither an immersive setting nor FL were the subject of investigation (n=380). After the first level screening 174 paper remained for further analyses.

STEP 3: eligibility

The exclusion criteria under Level 2 applied to the entire text. We excluded papers in which the study design did not include a qualitative, quantitative, or mixed-method measure. Additionally, studies were excluded for having no intervention described as immersive or without a focus on second/foreign

language learning, or acquisition. 35 of the identified papers did not present a study or study results. 36 studies did not examine immersive intervention or did not describe it as such. Another 34 papers did not conduct research in FL learning and teaching, and we did not have access to eight papers. After the second level screening 61 papers remained for further analysis.

STEP 4: final inclusion

Among the 61 identified papers, another seven papers were excluded due to lack of soundness of the findings and inconsistencies in the results. Finally, 54 papers were included for the systematic review (Figure 1).

4.2. Data Analysis

The following categories have further classified the 54 papers that fulfill the inclusion criteria:

Category 1: Investigation form and education level To identify research gaps and areas about target groups, the population participating in the empirical study was recorded. First of all, a distinction was made between teachers and learners, among the learners, regarding their education level (type of school, (college) student, or pupil), age, or proficiency level. Furthermore, a distinction was made between blended learning (BL) formats, experiments, and qualitative forms of inquiry. The main goal of BL design is to find the most effective and efficient combination of two learning modalities [71]. In this case, we distinguished between whether the immersive technology and the corresponding intervention were integrated into an experiment or into a several-weeks-long teaching concept.

Category 2: Degree of immersion and technology used As mentioned earlier, there are discussions in HCI around the definition of presence and immersion. Because authors used different definitions, we sorted the included articles by the definition used. As described above, defining immersion by technological factors, immersive applications can be categorized by the degree of immersion. The technologies used can be sorted by the degree of immersion. VR refers to the most immersive, followed by 360-degree formats shown via an HMD and AR. 360° formats which also include so-called CAVE systems or SmartSpaces. MUVE is the umbrella term for all interventions that do not use any technical means other than a mobile-, tablet, or desktop-based application, such as Second Life. In sum, this category reports which technological medium was used in the studies.

Category 3: Predictor In any empirical study, there is at least one independent variable that is manipulated. We would like to list all variables used in the identified studies to identify successful gestalt principles and possible research gaps and areas here as well.

Category 4: Criterion Similarly, we categorized the dependent or outcome variables to determine possible research areas and pinpoint outcomes with promising results. Additionally, this category provides an overview of the effects of using immersive interventions in second language learning and the closely linked intercultural learning.

5. Results

We examined the studies according to the four categories defined above.

5.1. Category 1: Investigation form and education level

A clear trend can be seen in the studies selected for the literature analysis. The ratio of pupils ($n = 25$ studies) compared to students ($n = 29$ studies) is almost equal. Teachers ($n = 8$) and others ($n = 2$) have been addressed less. Although eight studies interviewed teachers, most of these were observational or qualitative. Only in one study were educators, in this case university professors, the exclusive target group [62]. As can be seen in Figure 2, a distinction was made between blended learning (BL), experiment (between-design), experiment (within-design), and solely qualitative studies, such as semi-structured interviews. As the target groups in the studies were also mixed and different forms of investigation were used in one paper, the target groups summed up are larger in Figure 2.

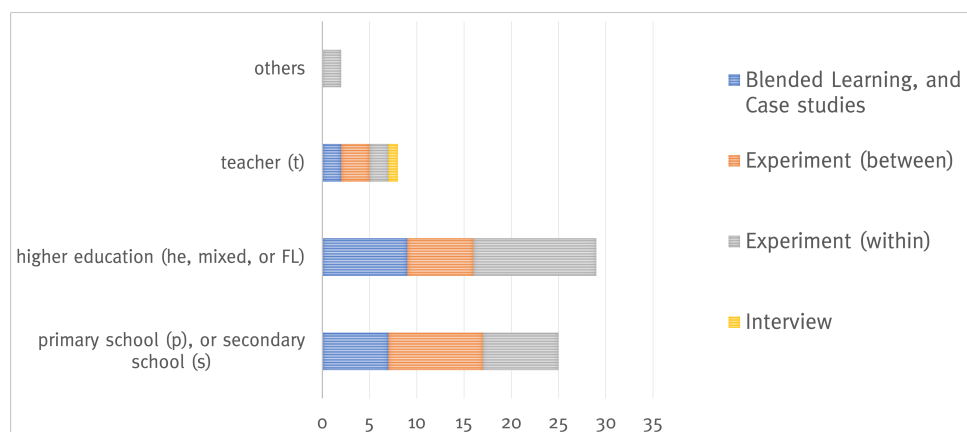


Figure 2. The table shows the distribution of the investigation form in relation to the education level of the participants.

The period in which the BL format was used in conjunction with an immersive intervention proceeded for up to one year. The results revealed more experiments ($n = 37$) than BL concepts ($n = 16$). We assume that it is due to the complexity of creating and integrating a BL design. The high number of studies with students (higher education) could be explained by the fact that they are better attracted to studies by credit programs, financial compensation, or similar. A total of four studies surveyed a mixed population [25, 45, 47, 51]. The analysis of investigation forms and the target group was studied to identify potential research gaps. Overall, blended learning concepts and the target group of teachers have rarely been addressed in the context of language and intercultural learning.

5.2. Category 2: Degree of immersion and technology used

Half of the studies used the AR medium (50%). The next largest share is taken by studies with MUVE's (24%). 13% of the studies showed the 360° content via an HMD or by Smart Spaces. Only in 13% of the studies fully immersive VR applications were deployed, and still none of these studies defined immersion according to the designated definition by Slater, and Wilbur [84]. In the few studies in which immersion was measured, either Witmer and Singer's definition [98] was used [94], or immersion was part of a larger questionnaire, such as the TAM questionnaire [65] or the game

experience questionnaire (GEQ) [28]. Yeh et al. (2020) used immersion as the quality of the learning VR system and surveyed it qualitatively by asking participants the following question, “Which of the following features (VR panorama, audio, interactive features, and structure) best help immersion in the VR experience?” The open-ended questions were evaluated using context analysis [102]. 19 Studies made no claims about immersion but used immersive technology by definition, namely AR. In 18 studies, “immersive” was an attribute of the application but no definition was given, and only eleven of these studies used AR, VR, or 360° content. However, the term “immersive”, or “immersion” has also been used to describe diving into the intervention [76, 93], in context with flow immersion [62], and cultural immersion [19, 82]. Quintín et al., for instance, referred to playing Second Life as “Immersion in Second Life” [76]. Immersion is most often defined in the field of HCI by Witmer and Singer, or by Slater and Wilbur’s terms of reference [84, 98]. Contrary to expectations, these definitions were rarely used in the identified studies. Only Liaw [63] and Wang et al. [94] used Witmer and Singer’s definition, and only Qu et al. [75] extended Witmer and Singer’s IPQ questionnaire [98] with the “one-item” regarding presence from Slater [88]. Gelsomini et al. [35] cited Johnson-Glenberg, et al. [53] as stating that immersion is a design consideration, and Guillen-Nieto et al. [38] cited de Freitas [27] regarding immersion. The definition of Shermann and Craig [80] was used by three studies [15, 49, 52]. Draxler et al. [29] used Sanchez-Vives and Slater’s [87] definition. It should be kept in mind that not all authors have the same basis and the same standards for immersion and presence. Thus, this makes the studies more difficult to compare in terms of their degree of immersivity. Therefore, results revealed a gap of interdisciplinary cross-discussion. While in the field of HCI, clear definitions of immersion and presence exist, including theories, measures, and a large corpus of empirical results. In the field of education, immersive technologies are used without such a clear definition and theoretical base.

5.3. Category 3: Predictor

Most studies compared the findings over time (31%), particularly in blended learning studies or studies with at least two measurement points. The second most common predictor was “degree of immersion”, where a traditional intervention was compared to an immersive intervention (28%). Studies with only one measurement time point (post measurements) consisted of 13% of the papers.

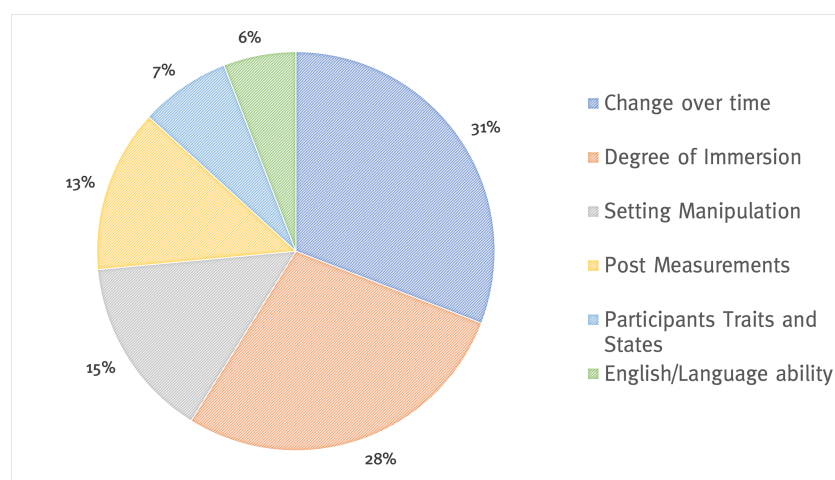


Figure 3. This pie-chart shows the reviewed predictors from the studies divided into six categories, and their frequency (in %).

5.3.1. Setting manipulation as predictor variable

A total of ten studies (15%) manipulated the setting of the application. Hoa, et al., tried to predict the learning effectiveness and motivation by different game designs (puzzle game, word matching, quiz, role-playing) [41], as well as, Hsu et al. [48]. Hsu, et al., used an AR game mechanism called collective game design (CGB), and another called sequential mission game design (SMG). *“In both systems, there were a total of seven reality targets corresponding to seven objects in the real-world setting. These seven objects could be collected randomly at the same ward in the CGB mode, while they had to be collected step by step at different wards in the SMG mode”* [48, p. 317]. While the teachers in the interview found that students in the CGB mode were significantly more enthusiastic, they made significant progress in their learning in the SMG mode (Paired sample $t = 2.34^*$; $p < 0.05$). Students in CGB mode did not make such remarkable progress (Paired sample $t = 1.46$; $p > 0.05$). The SMG mode forced students to study the learning material repeatedly until they passed the spelling task using the learning objective, which apparently helped them to remember the spelling [48]. In the study by Wang et al. (2017) there were four different intervention designs with two key artifacts. The first intervention was an English learning environment in a 3D virtual world, OpenSimulator, without any other artifacts. The second intervention included a chat-bot, the third a time machine, and the fourth the chat-bot and the time machine [94]. In summary, the chatbot and the time machine increased learners' sense of presence in a virtual environment. Learners who experienced the chatbot and those who experienced the time machine felt that they gained a sense of presence within the virtual environment. However, no data were provided here on the impact on FL learning.

Other studies have investigated whether the availability of certain support options makes a difference in language learning, such as scaffolding assistance (yes, no) [21], caption type (no caption, English caption, Chinese caption) [17] speech support (auditive, visual, or both) [92], speech input (yes, no) [26], or corrective feedback [25]. Liu et al. [66] investigated whether the social context during the use of an AR 3D pop-up book affects learning achievements. There were three trial settings: a) individual trial, b) small group trial, and c) real class trials. However, the setting's actual manipulations were examined and the participants' perceived effectiveness of the setting. The authors Chang et al. [12] investigated the effectiveness of the system using the TAM-questionnaire (Technology Acceptance Model) [65] which aimed to gather information on students' learning experience and attitude. Thus, they investigated whether the perceived effectiveness of the system influences behavioral intention, motivation, and satisfaction in learning English with an AR application [12].

5.3.2. Participants' traits and states as predictor variable

Things the participant already had as an ability or characteristic were chosen as predictor in 13% of the reviewed studies.

English or language ability was considered separately for this review, and accounted for 6% overall. Participants were divided into either two or three groups based on their scores on an English test or their school performance. Hence Gelsomini, et al. explored the question of whether children with more significant learning difficulties found it easier to remember vocabularies, and terms within IMAGINE [35].

Five studies (7%) used other traits or states of participants as predictors. For example, Hsu et al. [48] used gender as an independent variable. There was no remarkable influence on learning effectiveness due to the participants' gender. The authors of this study, therefore, recommend that other personal characteristics such as learning style or cognitive traits (working memory capacity, inductive

reasoning, and associative learning abilities) be considered in future studies [48]. Another study by Qu et al. divided the participants according to the attitude of their virtual counterparts. They used virtual bystanders in a virtual classroom setting to influence people's beliefs, self-efficacy, and anxiety [75]. In two other studies, the participants were divided according to their cognitive and learning styles [18,45]. In the study by Chang et al. [12], the user experience was the central object of investigation. On the one hand, the subjects' self-efficacy was used as a predictor. They found that perceived self-efficacy was a predictor of perceived learner satisfaction. The results also showed that perceived usefulness and perceived satisfaction were predictors of learners' behavioral intention to use e-learning. In addition, the authors of this study categorized the participants by their perceived effectiveness of the system itself. Therefore, the predictors of this study belong to the category of traits and states as well as to the category setting manipulation. A complete listing of the distribution of the predictors is shown in Figure 5.

As predictors, most studies took the most apparent change over time, and compared an immersive intervention to a traditional one. Of special interest were the results on participant characteristics and states as predictors, as well as setting manipulations. Specifically, gamification has a positive effect on motivation to learn, whereby designs with more scaffolding show better learning outcomes. Regarding states and traits, personal characteristics such as learning style or cognitive traits (working memory capacity, inductive reasoning, and associative learning skills) are recommended.

5.4. Category 4: Criterion

The following tables show the identified dependent variables sorted according to qualitative, and behavioral methods (see Tables 2, and 3), indicating the investigation form (questionnaire, eye-tracking data, etc.). Furthermore, the quantitative measured dependent variables are listed below (Table 4). Overall, learning achievements (examined in 31 studies) and qualitative and subjectively observed measures (examined in 35 studies) accounted for the largest proportions (Figure 4).

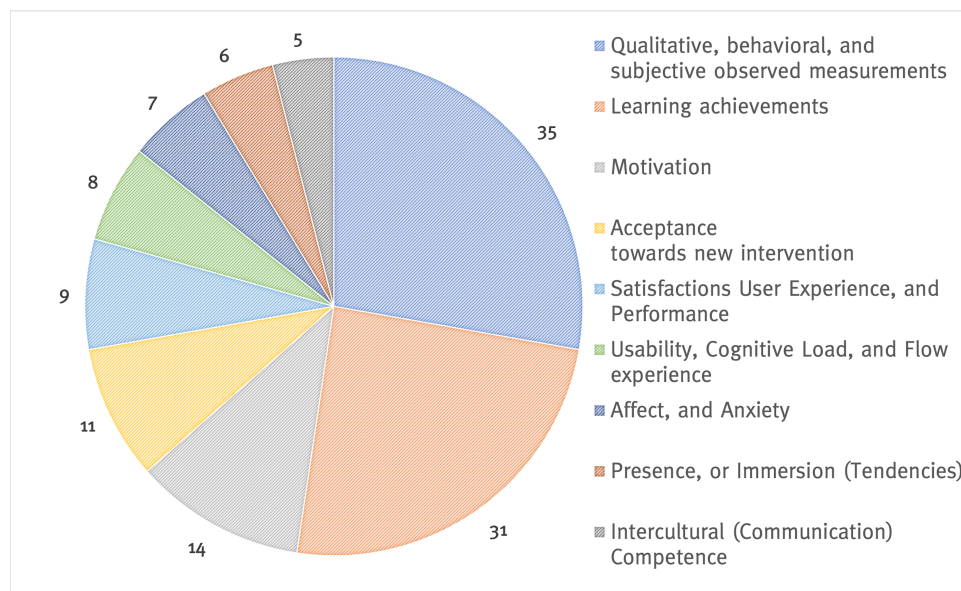


Figure 4. This pie chart shows the criteria, or responding dependent variables, that were examined in the studies reviewed. These were grouped into nine superordinate categories.

5.5. Results of qualitative measurements

Qualitatively, teachers often found the systems used very motivating and enjoyable [19,43,66]. They found it easy to use the systems, but only if pre-training had taken place [66]. Li et al. [62] conducted an exclusively qualitative study and interviewed six professors from the departments of Applied Linguistics and Linguistics, Information Management, and Information and Computer Engineering in semi-structured in-depth interviews after they were shown an AR learning environment. Although some technical problems occurred during the demonstration that caused delays, almost all of them perceived the AR system to be interesting, innovative, and attractive [62]. Some teachers assumed that AR might help students memorize vocabulary [66], while others perceived little vocabulary achievement [43]. Also frequently cited by teachers was the fact that application would distract users from the task because of its novelty. Instructors also mentioned that AR has an advantage over VR, due to greater peer communication and lower cost. Students and learners also increasingly mentioned in interviews the positive effect of interactivity, collaboration, and the resulting enjoyment and motivation from the new intervention. In none of the qualitative studies were learners found to have difficulties with the new technology.

Table 2. This table shows the studies that used qualitative measures and what type of investigation form was used.

number	study	qualitative investigation
23	[14, 62, 66]	Semi-Structured In-depth Interview.
	[36, 51]	Qualitative Feedback
	[47]	Group interview students, and teachers' comments
	[43]	Summary of teachers' comments
	[102]	Two open-ended questions
	[19, 48]	(Semi-structured) interviews with learners, and teachers
	[25, 76, 101]	Interview
	[50, 64]	Focus Group Interview
	[59]	Answers from video and log data
	[60, 93]	Reflection papers, or blogs
	[81, 91]	Ethnographic procedure, researcher as participant-observer; interviews
	[38]	Qualitative analysis of comments (positive, or negative)
	[58]	Qualitative data mentioned, but no further information.
[42]	Analysis of participants' speech	
[75]	Teachers' performance	

5.6. Results of observational, behavioral, and quantitative measurements

Behavioral measures confirmed the results on enjoyment and motivation. Experiential motivation was identified, for example, by longer game playing time, or more interaction, and enjoyment by increased, or extended, smiles and laughter [26]. There were also other observational and behavioral assessments. For example, Chen, and Hwang [18] analyzed the move structure during an oral presentation according to five main points and eleven sub-criteria. It was found that participants learning with the proposed ISVVR learning mode (360° content) used more move structures similar

to those of professional TED speakers compared to the conventional multimedia mode. The use of the immersive application (ISVVR) as a visual aid was able to improve the participants' movement structures in their oral presentations [18]. Ibrahim and colleagues investigated participants' attention during a conventional multimedia application (flashcard application) and the more immersive AR application (ARbis pictus) in terms of vocabulary learning behavior, recall, and recognition. For this purpose, gaze data were collected for both modalities. For the Flash Card application, eye-tracking data has been collected with a screen-based eye-tracker, and for the HoloLens application, head orientation focus has been recorded. Furthermore, click behavior on the multimedia application has also been collected to identify possible learning patterns. The differences in attention times between the different groups were not significant here. The click patterns showed that participants tended to click more frequently toward the end of the study. Users later reported in the interview that they found the ability to see the object and the word simultaneously in AR very helpful for learning. However, for autonomous self-testing, the flashcard application was more effective [51].

Physiological measures, for instance heart rate and skin conductance, were included to measure elicited arousal during the virtual English lesson (Table 3). Qu, et al.'s results [75] showed a significant effect for phasing on participants' heart rate, with an increase in participants' heart rate during the second phase in which they answered questions. In addition, the experience of witnessing bystanders commenting negatively on other students' performance increased participants' heart rates when it was their turn to speak.

Table 3. This table shows the studies that used behavioral, observational, subjective, or physiological measures and what type of investigation form was used.

number	study	behavioral investigation
12	[51]	attention and gaze (eye-tracking data)
	[3, 26, 29, 36, 38]	behavioral cues (subjective, observed data)
	[18]	move structures
	[25]	gaming behavior (log data)
	[50, 59, 76]	interaction patterns between students (video, and subjective, observed data)
	[75]	physiologically, through skin conductance and heart rate

5.6.1. Effects on language learning

Quantitatively, learning achievements, especially on a cognitive or affective level, were most frequently investigated, as well as the motivation and acceptance to use the new systems. Chen [20] reported in his study that the implemented method (ARVEL) effectively improves students' learning achievements and motivation. Students show a higher level of satisfaction with English as Foreign Language (EFL) learning than with the video-assisted learning method (CVEL). The ARVEL method significantly increased students' intrinsic motivation to learn English, and their satisfaction was significantly higher than the computer-based method. Wu [100] showed promising results in vocabulary learning success by comparing Pokémon Go with a flashcards application. Similarly, Redondo et al. [78] confirmed that the use of AR in EFL instruction in early childhood education produced higher levels of motivation in students than the more traditional instructional approach in the control group. Also, an improvement in English learning was observed in the children who used

AR in the classroom compared to those who used the traditional method. Students who practiced in IMAGINE showed that they retained many more concepts over the long term than those who practiced in the classroom. Children with more significant learning difficulties found it easier to remember terms within IMAGINE [35]. To study learning achievement in younger groups, matching games were used instead of tests, for example. Through AR, an improved classification of the vocabulary word and the corresponding picture, for instance, was determined here [3]. There were also studies comparing immersive interventions. Chen et al. [17], for example, compared interactions between subtitles and English proficiency in terms of motivation, and attitude, and effectiveness to learn in augmented-reality-enhanced theme-based contextualized EFL learning. Although all learners expressed positive attitudes toward AR-assisted learning, interactions between predictors (subtitles and English proficiency) and attitude were not significant. The captioning condition poorly influenced learning achievement.

5.6.2. Effects on intercultural learning

Intercultural competence was the least frequently used research item in the reviewed studies. In the study by Liaw [63], groups of different sizes (2-5 participants) from different parts of the world collaborated in vTime, a VR social networking site. Then, Intercultural Communication Competence (ICCQ, [68]) was assessed via questionnaire. The ICCQ results show that 88% of the participants were satisfied with the VR approach. 48% of participants disagreed that it was challenging to build a good relationship with people from other countries. Most participants (59%) also did not find dealing with and managing cultural uncertainties troublesome or frustrating (48%). Most of them found it quite enjoyable (60%) and even exciting (76%). Participants indicated that they had no prejudice against people who spoke with an accent (77%) and actively listened to others when they were in vTime interactions. All felt that they were attentive to the cultural and behavioral norms of others. [63].

Yeh and colleagues took a different approach [102]. Here, participants were asked to create their own VR content that presented the cultural significance of local Taiwanese sites in an educational VR application. The BL approach took 18 weeks to complete. In the first eleven weeks, the instructor taught participants how to use the EduVenture VR app to create panoramas and add interactive features. In weeks 12 through 15, participants created their content with the instructor's assistance by adding interactive features, panoramas, and audio. In the last three weeks, participants reviewed their classmates' work on the app. Then students were asked about their reactions to the experience with the VR cultural learning project. Immersion was not measured, but two open-ended questions were designed to explore how VR technology enriched students' intracultural learning experience (presence). Most students agreed that VR technology provided them with better intracultural learning experiences because VR technology made them feel like they were physically visiting and traveling to the places presented. All the features of VR helped students to present the local cultural sites in a vivid, attractive, and clear way [102].

Shih examined cultural knowledge acquisition and the emergence of attitudes toward the target culture [82]. She examined four participants in a case study for one year. They walked the virtual streets of London (Google Street View integrated into a virtual environment) with avatars' help under a native English teacher's guidance and interacted with each other via text and voice chat features. The results show that learners increased their cultural knowledge to varying degrees concerning culture's two main aspects - visible aspects and invisible aspects. There is a tendency for the effect of cultural immersion on participants' cultural knowledge acquisition to increase over time. All learners benefited from

cultural learning and showed positive attitudes toward the target culture, country, and interaction with English speakers to varying degrees after completing the course. The results suggest the possibility of a relationship between language level, character traits, motivation to learn about the target culture, initial attitudes, and cultural learning in terms of developing cultural knowledge and positive attitudes toward the target culture. Yang, and Liao investigated whether the learning effectiveness in terms of cultural learning can be increased by using an AR application (VECAR). This intervention was compared to Google Earth, which was operated via a screen by participants using the mouse. The results show that the effectiveness of learning cultural content was better in the experimental group using VECAR than in the control group. “*VECAR enables students to translate, rotate, scale, and modify virtual objects in 3D by using the intuitive hand gestures, thus it helps students to visualize and play around some cultural contents, such as football rules, that are difficult to understand solely based on verbal explanation*” [101, p. 114].

The last of the five studies that investigated intercultural competence examined the language proficiency of the participants. Here, they played a serious game (It’s a deal) for teaching intercultural business communication. The statistical test for the correlation between learning effectiveness and English proficiency level showed that the students of A2, B1, and B2 proficiency levels made more intercultural learning progress than the students of C1 and C2 proficiency levels. Accordingly, the native speakers participating in the surveyed had the least learning progress [38]. However, the limitation of this study is the low immersion. The serious game was played through a computer. Immersion has been described as part of a dimension in a framework for educational games [27].

5.7. Combination, and summary of the results

The explanatory variable in over half of the studies was change over time, or degree of immersion. As shown in the following plot, no study examined how immersive their application was when compared to a traditional intervention. The following figure shows the predictors in combination with the criteria. All pairs of variables that appeared in the identified studies were combined here. Because not every study examined only one independent and one dependent variable, there are more combinations than studies together (see also Table 4).

Most frequently, subjects’ learning achievements were examined in combination with the degree of immersion (immersive vs. traditional setting). This combination was investigated in a total of 16 studies. Often the subject of the investigation was learning of vocabulary. Most studies found improved language skills. However, researchers considered that questionnaires and achievement tests might not be the best way to obtain information from children as they are still young to read and understand [46]. Also, very often, the learning achievements over time were examined (N=14), as well as the motivation or the acceptance towards the new technology used predicted by the degree of immersion was examined (N=12). The motivation was usually higher in the more immersive interventions, and attitudes and acceptance of the new technology were also high. The use (usability, satisfaction, and user experience) was easy for most participants. However, no study investigated why immersive technologies brought better results and which gestalt features have been successful. As stated above, studies addressing a different focus already revealed some specific characteristics of immersive technologies. Future studies might investigate whether those characteristics are useful gestalt features for language and intercultural learning. As already mentioned, fully, immersive technologies, such as VR, provide opportunities for perspective-taking, intercultural encounters, and changes in certain attitudes and behaviors. Nevertheless, the results of this review show that few studies

used fully immersive interventions (N=7).

Moreover, in only eight studies were teachers the participants (Table 4). They mostly took an observational role. Only one study surveyed solely teachers qualitatively [62], and only one other study investigated teachers' communication behavior quantitatively [101].

Zhang, and Zhou's [103] assertion that language skills are increasingly investigated rather than intercultural competences or conative objects of inquiry, can be confirmed. Nowadays, intercultural competence is a learning objective that should be taken into account when teaching foreign languages. Nevertheless, it has been investigated in the reviewed studies only by five studies and thus the least frequently (Table 4). Constructive views on language learning and teaching focus additionally on intercultural competencies [33]. Communication in realistic contexts was identified as one primary method to improve knowledge about the language itself and the culture and traditions. Although intercultural learning is an essential part of modern language learning and immersive technologies are proven to be useful in this context, only a few reviewed studies addressed this topic. Thus, future studies might focus on intercultural aspects and ask how immersive technologies can increase intercultural learning in language learning and teaching.

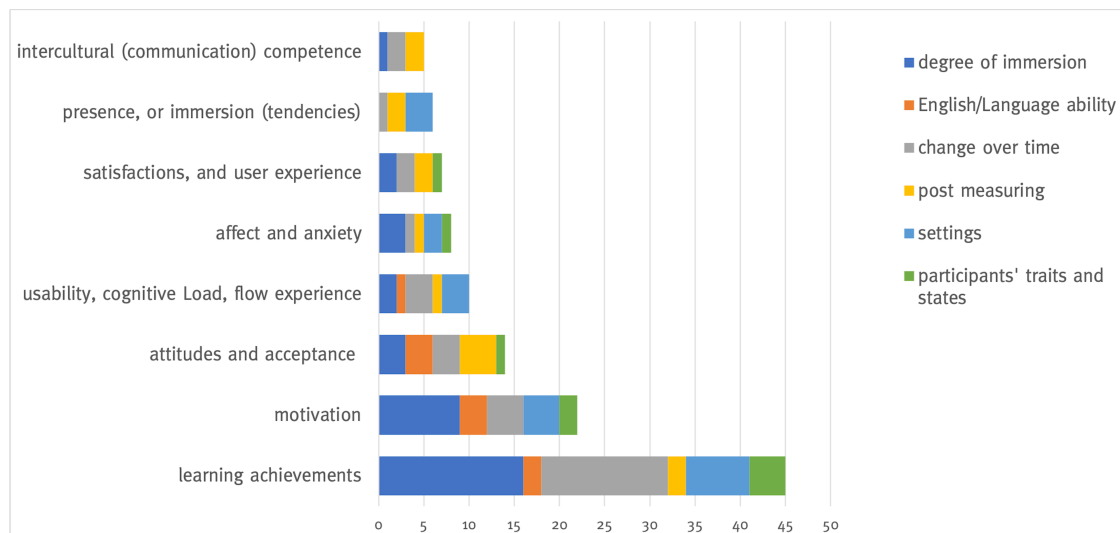


Figure 5. This plot shows the existing combination pairs of predictors and criterions of the reviewed studies and their frequencies.

Table 4. This table summarizes the study results from all studies that used at least one quantitative measurement. These are sorted by the previously defined categories of (1) degree of immersion (2) investigation form, and education level (3) predictor's, and (4) criterion's.

application	immersion		education level			investigation form	predictor's	criterion's
	p	s	he (mixed)	he (FL)	teacher			
MARAEALS [46] POSTECH [59]	✓	✓				Blended learning Blended learning	post evaluation English ability	acceptance Table 2, and 3 acceptance culture learning experience, Table 2
EduVenture VR [102]			✓			Blended learning	post evaluation	learning achievements, Table 2
Second Life [93]			✓			Blended learning	change over time degree of immersion	acceptance usability satisfaction, performance, and user experience
OpenSim [67]			✓			Blended learning	change over time	anxiety, immersion usability, task load, satisfaction
CIR [11]			✓		✓	Blended learning	change over time	learning achievements, acceptance, Table 2
Google Earth VR [19]						Blended learning	change over time	Table 2
Second Life [64] AR-application [66] VEC3D [81]		✓		✓	✓	Blended learning Blended learning Blended learning	change over time post evaluation change over time	Table 2 Table 2 Table 2 satisfaction and performance
vTime [63]			✓			Blended learning	post evaluation	presence intercultural communication competence
AR-application [56]		✓				Blended learning	English ability change over time	acceptance, cognitive load learning achievements motivation anxiety Table 3
ISVVR [18]			✓			Quasi experiment	degree of immersion traits, and states	learning achievements motivation anxiety Table 3
AR 3D book [91]	✓					Quasi experiment	change over time setting manipulation	learning achievements, Table 2
Mobile-AR [14]	✓					Quasi experiment	English ability traits and states	learning achievements motivation Table 2
Google Street View [82]				✓		Case study	change over time	learning achievements cultural knowledge

application	immersion	education level		investigation form	predictor's	criterion's
		he (mixed)	he (FL)			
ARVEL [20]	AR			Experiment(between design)	degree of immersion	learning achievements motivation, satisfaction, and performance
Aurasma, and Quiver [78]	AR			Experiment(between design)	degree of immersion	motivation (enjoyment) affect
Mobile-AR [43]	AR		✓	Experiment(between design)	degree of immersion	learning achievements, Table 2
MOW [3]	AR			Experiment(between design)	degree of immersion	learning achievements usability Table 3
Interactive Tablet-Game [50]	MUVE			Experiment (between design)	degree of immersion English ability	motivation acceptance interaction patterns Table 2, and 3
Mobile-AR [16]	AR			Experiment (between design)	degree of immersion change over time	learning achievements motivation
Serious VR-Game [54]	VR			Experiment (between design)	degree of immersion change over time	learning achievements motivation
Mobile-AR [41]	AR			Experiment (between design)	degree of immersion setting manipulation	learning achievements motivation
VR-application [52]	360°	✓		Experiment (between design)	change over time	learning achievements motivation
Pokémon Go [100]	AR	✓		Experiment(between design)	degree of immersion	learning achievements acceptance
Mobile-AR [21]	AR	✓		Experiment (between design)	setting manipulation	satisfaction, learning achievements
VILLAGE [94]	MUVE	✓		Experiment(between design)	setting manipulation	learning achievements usability
Second life [23]	MUVE	✓		Experiment(between design)	degree of immersion	presence, and immersion tendencies learning achievements motivation
Mobile-AR [29]	AR	✓		Experiment (between design)	change over time degree of immersion	acceptance learning achievements Tables 3
VECAR [101]	AR	✓		Experiment (between design)	degree of immersion	usability cultural learning effectiveness, affect Table 2
Mobile-AR [48]	AR	✓		Experiment (between design)	setting manipulation traits and states	learning achievements cognitive load flow experience Table 2
Mobile-AR [17]	AR	✓		Experiment (between design)	English ability setting manipulation	motivation, learning achievements

application	immersion	p	s	education level		investigation form	predictor's	critereon's
				he (mixed)	he (FL)			
VR-application [42]	VR			✓		Experiment(within design)	change over time	learning achievements, Table 2
IMAGINE [35]	360°	✓				Experiment (within design)	degree of immersion change over time	learning achievements
TeachAR [26]	AR	✓				Experiment(within design)	degree of immersion setting manipulation	learning achievements Table 3
Second Life [58]	MUVE	✓				Experiment (within design)	change over time	learning achievements Table 2)
VR-application [49]	VR	✓				Experiment(within design)	post evaluation	acceptance
Mobile-AR [5]	AR	✓				Experiment (within design)	post evaluation	learning achievements
Location-based AR [92]	AR	✓				Experiment (within design)	change over time setting manipulation	learning achievements motivation
VR-application [15]	VR			✓		Experiment (within design)	degree of immersion change over time	learning achievements
It's a Deal! [38]	MUVE				✓	Experiment(within design)	change over time	intercultural awareness, and knowledge Table 2, and 3
Second Life [22]	MUVE				✓	Experiment (within design)	degree of immersion	learning achievements motivation
VEDILS [6]	360°			✓		Experiment (within design)	post evaluation	acceptance, satisfaction, and performance, usability, anxiety, immersion
Marker-based AR [12]	AR			✓		Experiment(within design)	traits and states	satisfaction, and performance acceptance
Second Life [76]	MUVE			✓		Experiment (within design)	change over time	learning achievements Table 2, and 3
VR environment [75]	VR			✓		Experiment(within design)	setting manipulation	satisfaction, performance, anxiety, presence, Table 2, and 3
Location-based AR [60]	AR			✓		Experiment (within design)	post evaluation	learning achievements Table 2
ARELS [47]	AR			✓		Experiment(within design)	change over time	motivation Table 2
Arbis pictus [51]	AR			✓		Experiment(within design)	degree of immersion	learning achievements Table 2, and 3
Wikitude World Browser with AR [45]	AR			✓		Experiment(within design)	traits and states	learning achievements
CARDS [36]	AR		✓			Experiment (within design)	post evaluation	user experience Table 2, and 3
3D game [25]	MUVE		✓			Experiment(within design)	setting manipulation	affect immersion motivation Table 2, and 3

6. Discussion

At the onset, the following questions were posed, which are to be answered with the newly gained knowledge:

RQ1: How are virtual, fully immersive learning environments used for foreign language learning?

It was observed that a large number of studies compare conventional teaching methods with interventions described as immersive for foreign language learning. The criteria studied were mostly cognitive learning achievements, such as vocabulary or speaking skills, or affective variables, such as motivation, satisfaction or speaking anxiety and discomfort. Measurements on a conative level are quite rare. Only in two studies, the behavior was surveyed by collecting eye-tracking data and by using physiological measurement methods. We see great potential here, as these measurements are almost uncontrollable by the participants.

We also see great potential and a present research gap in the area of intercultural and transcultural language learning. Intercultural competence was the subject of research in only five identified studies. As mentioned at the beginning, Chen, and Starosta [13] assume that three levels define intercultural competence. At the cognitive level, intercultural awareness should be created, intercultural sensitivity should be addressed at the affective level, and the conative level should address intercultural fluency [13, 103]. Current research provides little on the topic of intercultural competence nor at the cognitive, or affective neither the conative level. Intercultural sensitivity is less studied than criteria such as motivation or speech anxiety. In today's English classes, teachers, and learners have to acquire knowledge, develop a critical, and sensitive awareness of racial thinking and act with respect and open-mindedness towards one's own and others' perspectives [30, 31]. Particularly with the renewed attention to the BlackLiveMatters movement, we see that this learning objective is of enormous importance. It should therefore be given a corresponding priority in language education. Fully immersive learning environments have been used in most of the studies identified for foreign language learning by examining primarily cognitive teaching, and learning processes, such as vocabulary learning achievements, or affective investigations, such as motivation or speaking anxiety. Constructive views of language learning and teaching additionally focus on intercultural competencies [33]. Intercultural learning is an essential component of modern language learning. In the context of immersive technologies, few reviewed studies have addressed this topic. Researchers should therefore focus on intercultural aspects in the future and ask how immersive technologies can enhance intercultural learning in language learning and teaching.

RQ2: Which characteristics of immersive technology support language and intercultural learning?

The benefit of increased immersion was mentioned in most studies but not checked. Only six studies quantitatively measured immersion or presence. In the three studies that measured immersion, it was only one item of a broader questionnaire. In one study, the presence was measured qualitatively [102]. Hence, studies addressing language and intercultural learning with immersive technologies lack systematic investigations on the increment value of immersive technologies and different immersion degrees. Furthermore, as can be seen in Figure 5, no study comparing an immersive intervention with a traditional intervention measured immersion or presence. In other words, no study tested whether the interventions were truly distinct enough in their immersiveness for further assumptions. Moreover, only seven of the 54 papers offered fully immersive settings. Thus, the spectrum of

immersive technologies is unexplored - particularly full immersive technologies. Future researchers should therefore focus on the area of foreign language and intercultural learning and teaching using fully, immersive VR applications. Liaw's study [63] also shows that VR can offer great potential for use in the classroom and especially for intercultural encounters (41% of participants felt that VR helped them learn). Yet, the other 41% did not seem to feel such effects, and 18% responded negatively to this statement. Only 18% of the participants felt that using VR fit their learning style. Therefore, it is imperative to focus on interdisciplinary collaboration when developing fully immersive learning environments [63].

The Behavioral Framework for Immersive Technologies (BehaveFiT) presents four main factors that can be manipulated in VR environments to promote change in attitude and behavior [96]. The first factor is the situational context, i.e. the appearance of the virtual environment in which the user feels present. In the study by Yeh and colleagues, participants were allowed to create their own VR environment. Students ranked panoramic (situational context) as the most effective technology to deliver an immersive cultural learning experience, followed by audio, interaction, and structure [102]. The second factor is the virtual objects and their appearance, interactivity etc. As Yang and Liao showed, the AR application VECAR enabled students to translate, rotate, scale, and modify virtual objects in 3D using intuitive hand gestures. This allowed cultural content to be easily and quickly visualized and thus paraphrased [101]. Object integration in VR is almost limitless and can be integrated effectively into foreign language learning. Nevertheless, because of this limitlessness, especially with regard to cultural learning, it is essential to pay attention to possible associations and connotations of the objects in order to avoid critical incidents. The third and fourth factors are self-presentation and other-presentation. For example, in the study by Qu et al. [75], virtual bystanders perceived attitudes toward participants and virtual peers were influenced by their appearance. Bystanders' attitudes toward participants significantly influenced participants' performance beliefs, self-efficacy, or anxiety. Virtual bystanders who displayed consistently positive attitudes toward both peer speakers and participants elicited the least anxiety among participants. Furthermore, participants rated their presence higher when the bystanders exhibited a positive instead of a negative attitude towards them [75]. Thus some studies showed positive effects. However, future research on language and intercultural learning might draw on those promising research stemming from other fields (e.g., [1, 2, 44, 74, 77]).

RQ3: Can virtual, fully immersive learning environments increase motivation and success in learning a foreign language?

Immersive applications can increase motivation and are highly accepted by learners. There are already some promising research results in this area. At the same time, researchers, as well as teachers, warn against the strong flood of stimuli, the effect of the new, and the distracting effect of animated objects, especially on younger participants [46]. In terms of learning success, nearly all studies measured an improvement in learning achievements.

RQ4: Can virtual, fully immersive learning environments change participants' attitudes through intercultural encounters?

Affective behavioral changes in non-educational VR applications have been extensively studied [39]. However, they are underrepresented in educational settings and represent an important area for future research. There is only one study that demonstrates that virtual encounters impact our self-

efficacy, speaking anxiety, and avoidance behaviors in the foreign language classroom. However, it was not the cross-cultural component that was manipulated, but the virtual bystanders' attitudes toward the participants and the virtual peer speakers. Thus, there is still a large research gap that has not been explored, especially in the pedagogical framework and in the field of foreign language teaching, which, as the Covid-19 crisis shows, should receive special attention.

RQ5: How are virtual, fully immersive learning environments used for teacher training?

Teachers were rarely included in the evaluation and usually only took part as observers. This was then surveyed qualitatively in only seven studies. Just one study investigated the teacher communication behavior quantitatively as well as qualitatively [101]. Since student teachers are relevant future multipliers of teaching and learning with digital media in the future, they might be addressed twofold [97]. Firstly, the teaching and learning processes at the university set an example for the student teachers by imparting digital key competencies. There are some promising projects and suggestions on how to address student teachers. For example, the prospective teachers can explore and experience various digital devices, and can be informed about the basic functionalities. They can receive a reflected understanding of the benefits and limitations of digital devices and develop digital competencies by working on projects with digital media. Further, students can learn how to use digital devices to implement pedagogical as well as methodological objectives. Finally, the students can evaluate their material during the course, imparting knowledge about the appropriate use of digital equipment to attain didactical aims. Secondly, the prospective teachers prepare teaching units for practical implementation. Thus, innovative teaching and learning contexts can be directly explored in realistic situations and transferred into real-life teaching scenarios in schools [31, 79, 97].

7. Conclusion

Some benefits of immersive technology and its use in foreign language teaching were utilized and evidenced within the identified studies, such as increased attention, motivation, and enjoyment. Nevertheless, it also revealed research gaps concerning implicit measurements of learning achievements, and intercultural competence on a cognitive, affective, and conative level. In modern foreign language teaching, the focus is no longer exclusively on native language competence. It is now common to see the overall goal of foreign language teaching as the development of intercultural and language action-oriented communicative competence, which consciously includes the development of competencies in the learners' inter-language. *"This change in goals marks a shift away from a standards-based deficit orientation toward the speaker as a linguistically and culturally shaped individual. A language embodies and expresses aspects of its speakers' culture and worldview. Full communicative competence in a language, therefore, requires an understanding of and the ability to interact with the culture and worldview of the speakers of that language"* [7, p. 2]. Not only this, but also the renewed Black Lives Matter movement indicate that inter- and trans-cultural competencies are not given enough space in the foreign language classroom compared to deficit-oriented language acquisition measured by vocabulary and grammar. Yet, as this review shows, the study of intercultural competence through immersive technologies is a research gap that needs to be filled.

Furthermore, the usage of fully, immersive interventions, like VR, and its associated benefits, were comparatively rarely used. Especially in the context of studies and basic research, VR environments offer a comparatively cheap and very controllable setting. Studies are more reproducible and therefore

more comparable than conventional teaching methods. The systematic review further revealed that many potentials remained unexploited. For example, spatial distances and temporal distances can be overcome, which might support the learning of tempo and intercultural competence. Especially in times of pandemics, it can be a great advantage to be able to meet remotely at a virtual location. Simultaneously, the design of this location can be limitless and easily changed. Furthermore, the representation of the other persons, and the self-presentation can be manipulated. Studies from behavioral psychology show here promising results, which would be applicable particularly in the area of intercultural competence. However, stylized avatars can also have advantages. For example, anonymity can arise, which can have a positive impact on speech anxiety. In addition to these factors, virtual objects can be integrated, which might steer and inspire discussions. Visual feedback and learning progress can be displayed directly by an instructor, and made tangible, even without a student feeling exposed in front of a class. Future researchers on language and intercultural learning should therefore focus on these potentials of immersive technologies to design innovative and effective learning environments.

Immersive learning environments can address the challenges of digitization in educational systems. The first success factor of digitization is the digital equipment of schools. The second is acceptance among teachers and their digital skills in using the equipment. Although this is the case, the target group of educators was heavily underrepresented in this review. The third success factor of digitization is interdisciplinary cooperation. These seem essential in order to exploit the whole potential of digitized programs. Both pedagogical and didactic concepts must be combined here with research from HCI to achieve innovative teaching-learning formats and a successful digital transformation.

Acknowledgments

This research has been funded by the Federal Ministry of Education and Research (BMBF): “Funding for teacher training with a focus on digitalization in teacher training” in the project CoTeach (project number 01JA2020).

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. Ahn SJ, Bostick J, Ogle E, et al. (2016) Experiencing nature: Embodying animals in immersive virtual environments increases inclusion of nature in self and involvement with nature. *Journal of Computer-Mediated Communication* 21: 399–419.
2. Banakou D, Groten R, Slater M (2013) Illusory ownership of a virtual child body causes overestimation of object sizes and implicit attitude changes. *Proceedings of the National Academy of Sciences* 110: 12846–12851.
3. Barreira J, Bessa M, Pereira LC, et al. (2012) MOW: Augmented Reality game to learn words in different languages: Case study: Learning English names of animals in elementary school. In *7th Iberian conference on information systems and technologies (CISTI 2012)*. 1–6. IEEE.
4. Barrett MD (2011) Intercultural competence. *EWC Statement Series 2*: 23–27.

5. Bazzaza MW, Alzubaidi M, Zemerly MJ, et al. (2016) Impact of smart immersive mobile learning in language literacy education. In *2016 IEEE Global Engineering Education Conference (EDUCON)*. 443–447. IEEE.
6. Berns A, Mota JM, Ruiz-Rube I, et al. (2018) Exploring the potential of a 360 video application for foreign language learning. In *Proceedings of the Sixth International Conference on Technological Ecosystems for Enhancing Multiculturality*. 776–780.
7. Blell G, Doff S (2014) Mehrsprachigkeit und Mehrkulturalität: Einführung in das Thema. *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 19.
8. Bond M, Marín VI, Dolch C, et al. (2018) Digital transformation in German higher education: student and teacher perceptions and usage of digital media. *Int J Educ Technol H* 15: 1–20.
9. Buhl H, Winter R (2009) Full virtualization – BISE’s contribution to a vision. *Business and Information Systems Engineering* 1: 133–136.
10. Byram M, Doyé P (1999) Intercultural competence and foreign language learning in the primary school. *The teaching of modern foreign languages in the primary school*, 138–151.
11. Chabot S, Drozdal J, Peveler M, et al. (2020) A collaborative, immersive language learning environment using augmented panoramic imagery. In *2020 6th International Conference of the Immersive Learning Research Network (iLRN)*, 225–229.
12. Chang YJ, Chen CH, Huang WT, et al. (2011) Investigating students’ perceived satisfaction, behavioral intention, and effectiveness of English learning using augmented reality. In *2011 IEEE International Conference on Multimedia and Expo*. 1–6.
13. Chen G, Starosta W (1999) A review of the concept of intercultural awareness. *Human Communication* 2: 27–54.
14. Chen CP, Wang CH (2015) The effects of learning style on mobile augmented-reality-facilitated English vocabulary learning. In *2015 2nd International Conference on Information Science and Security (ICISS)*. 1–4.
15. Chen YL (2016) The effects of virtual reality learning environment on student cognitive and linguistic development. *The Asia-Pacific Education Researcher* 25: 637–646.
16. Chen SY, Hung CY, Chang YC, et al. (2018) A study on integrating augmented reality technology and game-based learning model to improve motivation and effectiveness of learning English vocabulary. In *2018 1st International Cognitive Cities Conference (IC3)*. 24–27.
17. Chen MP, Wang LC, Zou D, et al. (2020) Effects of captions and English proficiency on learning effectiveness, motivation and attitude in augmented-reality-enhanced theme-based contextualized EFL learning. *Computer Assisted Language Learning*, 1–31.
18. Chen MRA, Hwang GJ (2020) Effects of experiencing authentic contexts on English speaking performances, anxiety and motivation of EFL students with different cognitive styles. *Interactive Learning Environments*, 1–21.
19. Chen Y, Smith TJ, York CS, et al. (2020) Google Earth Virtual Reality and expository writing for young English Learners from a Funds of Knowledge perspective. *Computer Assisted Language Learning* 33: 1–25.
20. Chen CH (2020) AR videos as scaffolding to foster students’ learning achievements and motivation in EFL learning. *British Journal of Educational Technology* 51: 657–672.
21. Chew SW, Jhu JY, Chen NS (2018) The effect of learning English idioms using scaffolding strategy through situated learning supported by augmented reality. In *2018 IEEE 18th International Conference on Advanced Learning Technologies (ICALT)*. 390–394.

22. Chung LY (2011) Using avatars to enhance active learning: Integration of virtual reality tools into college English curriculum. In *The 16th North-East Asia Symposium on Nano, Information Technology and Reliability*. 29–33.
23. Chung LY (2012) Virtual Reality in college English curriculum: Case study of integrating second life in freshman English course. In *2012 26th International Conference on Advanced Information Networking and Applications Workshops*. 250–253.
24. Cooper C, Varley-Campbell J, Booth A, et al. (2018) Systematic review identifies six metrics and one method for assessing literature search effectiveness but no consensus on appropriate use. *J clin epidemiol* 99: 53–63.
25. Cornillie F, Clarebout G, Desmet P (2012) Between learning and playing? Exploring learners' perceptions of corrective feedback in an immersive game for English pragmatics. *ReCALL: Journal of Eurocall* 24: 257–278.
26. Dalim CSC, Dey A, Piumsomboon T, et al. (2016) TeachAR: An interactive augmented reality tool for teaching basic English to non-native children. In *2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*. 82–86.
27. De Freitas S (2006) Learning in immersive worlds: A review of game-based learning. 2–60.
28. De Grove F, Van Looy J, Courtois C (2010) Towards a serious game experience model: Validation, extension and adaptation of the GEQ for use in an educational context. In *Playability and player experience*. 10: 47–61. Breda University of Applied Sciences.
29. Draxler F, Labrie A, Schmidt A, et al. (2020) Augmented reality to enable users in learning case grammar from their real-world interactions. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–12.
30. Eisenmann M, Grimm N, Volkmann L (2010) 92. Teaching the new English cultures and literatures. *English and American Studies in German* 2009: 164–166.
31. Eisenmann M (2019) Teaching English: Differentiation and Individualisation. utb GmbH.
32. Fast-Berglund Å, Gong LL (2018) Testing and validating extended reality (xR) technologies in manufacturing. *Procedia Manuf* 25: 31–38.
33. Fernández SS, Pozzo MI (2017) Intercultural competence in synchronous communication between native and non-native speakers of Spanish. *Language Learning in Higher Education* 7: 109–135.
34. Garzon J, Acevedo J (2019) Meta-analysis of the impact of Augmented Reality on students' learning gains. *Educational Research Review* 27: 244–260.
35. Gelsomini M, Leonardi G, Garzotto F (2020) Embodied learning in immersive smart spaces. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–14.
36. Giraudeau P, Olry A, Roo JS, et al. (2019) CARDS: a mixed-reality system for collaborative learning at school. In *Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces*. 55–64.
37. Grünwald A (2016) Üben und Übungen im Fremdsprachenunterricht. *Üben und Übungen beim Fremdsprachenlernen: Perspektiven und Konzepte für Unterricht und Forschung. Arbeitspapiere der 36. Frühjahrskonferenz zur Erforschung des Fremdsprachenunterrichts*, 84.
38. Guillén-Nieto V, Aleson-Carbonell M (2012) Serious games and learning effectiveness: The case of It's a Deal! *Computers and Education* 58: 435–448.

39. Hamilton D, McKechnie J, Edgerton E (2020) Immersive virtual reality as a pedagogical tool in education: a systematic literature review of quantitative learning outcomes and experimental design. *Journal of Computer Education* 8: 1–32.
40. Hammer M (2012) The intercultural development inventory: A new frontier in assessment and development of intercultural competence. Sterling, VA: Stylus Publishing. In M. Vande Berg, R. M. Paige, and K. H. Lou (Eds.). *Student learning abroad*, 115–136.
41. Hao KC, Lee LC (2019) The development and evaluation of an educational game integrating augmented reality, ARCS model, and types of games for English experiment learning: an analysis of learning. *Interactive Learning Environments*, 1–14.
42. Hassani K, Nahvi A, Ahmadi A (2016) Design and implementation of an intelligent virtual environment for improving speaking and listening skills. *Interactive Learning Environments* 24: 252–271.
43. He J, Ren J, Zhu G, et al. (2014) Mobile-based AR application helps to promote EFL children's vocabulary study. In *2014 IEEE 14th International Conference on Advanced Learning Technologies*. 431–433.
44. Herrera F, Bailenson J, Weisz E, et al. (2018) Building long-term empathy: A large-scale comparison of traditional and virtual reality perspective-taking. *PLOS ONE* 13: e0204494.
45. Ho SC, Hsieh SW, Sun PC, et al. (2017) To activate English learning: Listen and speak in real life context with an AR featured u-learning system. *Journal of Educational Technology and Society* 20: 176–187.
46. Hsieh M (2016) Development and evaluation of a mobile AR assisted learning system for English learning. *2016 International Conference on Applied System Innovation (ICASI)*, Okinawa, 1-4.
47. Hsieh MC (2016) Teachers' and students' perceptions toward augmented reality materials. In *2016 5th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI)*. 1180–1181.
48. Hsu TC (2019) Effects of gender and different augmented reality learning systems on English vocabulary learning of elementary school students. *Universal Access in the Information Society* 18: 315–325.
49. Huang X, Han G, He J, et al. (2018) Design and Application of a VR English Learning Game Based on the APT Model. In *2018 Seventh International Conference of Educational Innovation through Technology (EITT)*. 68–72.
50. Hung HC, Young SSC (2015) An investigation of game-embedded handheld devices to enhance English learning. *Journal of Educational Computing Research* 52: 548–567.
51. Ibrahim A, Huynh B, Downey J, et al. (2018) Arbis pictus: A study of vocabulary learning with augmented reality. *IEEE T Vis Comput Gr* 24: 2867–2874.
52. Ji S, Li K, Zou L (2019) The Effect of 360-Degree Video Authentic Materials on EFL Learners' Listening Comprehension. In *2019 International Joint Conference on Information, Media and Engineering (IJCIME)* . 288–293.
53. Johnson-Glenberg MC, Birchfield DA, Tolentino L, et al. (2014) Collaborative embodied learning in mixed reality motion-capture environments: Two science studies. *Journal of Educational Psychology* 106: 86.
54. Khatoony S (2019) An Innovative Teaching with Serious Games through Virtual Reality Assisted Language Learning. In *2019 International Serious Games Symposium (ISGS)*. 100–108.

55. Kincheloe Joe L (2008) *Critical Pedagogy Primer 2nd edition*. English: New York, NY: Peter Lang.
56. Küçük S, Yılmaz RM, Göktaş Y (2014) Augmented reality for learning English: Achievement, attitude and cognitive load levels of students. *Education and Science/Eğitim ve Bilim* 39.
57. Küster L (2014) Zur Einführung in den Themenschwerpunkt. *Fremdsprachen lehren und lernen* 43: 2.
58. Lan YJ (2015) Contextual EFL learning in a 3D virtual environment. *Language Learning and Technology* 19: 16–31.
59. Lee K, Kweon SO, Lee S, et al. (2014) POSTECH immersive English study (POMY): Dialog-based language learning game. *IEICE T Inf Syst* 97: 1830–1841.
60. Lee SM, Park M (2020) Reconceptualization of the context in language learning with a location-based AR app. *Computer Assisted Language Learning* 33: 936–959.
61. Leyva F, Plummer CJ (2015) National Institute for Health and Care Excellence 2014 guidance on cardiac implantable electronic devices: health economics reloaded.
62. Li KC, Tsai CW, Chen CT, et al. (2015) The design of immersive English learning environment using augmented reality. In *2015 8th International Conference on Ubi-Media Computing (UMEDIA)*. 174-179.
63. Liaw ML (2019) EFL learners' intercultural communication in an open social virtual environment. *Journal of Educational Technology and Society* 22: 38–55.
64. Liou HC (2012) The roles of Second Life in a college computer-assisted language learning (CALL) course in Taiwan, ROC. *Computer Assisted Language Learning* 25: 365–382.
65. Liu IF, Chen MC, Sun YS, et al. (2010) Extending the TAM model to explore the factors that affect intention to use an online learning community. *Computers and education* 54: 600–610.
66. Liu E, Liu C, Yang Y, et al. (2018) Design and implementation of an augmented reality application with an English Learning Lesson. In *2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALe)* 494–499.
67. Lorenzo CM, Lezcano L, Alonso SS (2013) Language Learning in Educational Virtual Worlds-a TAM Based Assessment. *J UCS* 19: 1615–1637.
68. Matveev AV (2002) The perception of intercultural communication competence by American and Russian managers with experience on multicultural teams (Doctoral dissertation, Ohio University).
69. Milgram P, Kishino F (1994) A taxonomy of mixed reality visual displays. *IEICE T Inf Syst* 77: 1321–1329.
70. Moher D, Liberati A, Tetzlaff J, et al. (2011) Bevorzugte Report Items für systematische Übersichten und Meta-Analysen: Das PRISMA-Statement. *DMW-Deutsche Medizinische Wochenschrift* 136: e9–e15.
71. Neumeier P (2005) A closer look at blended learning—parameters for designing a blended learning environment for language teaching and learning. *ReCALL: the Journal of EUROCALL* 17: 163.
72. Oberdörfer S, Latoschik ME (2019) Predicting learning effects of computer games using the Gamified Knowledge Encoding Model. *Entertain Comput* 32: 100315.
73. Oberdörfer S, Elsässer A, Schraudt D, et al. (2020) Horst-The teaching frog: learning the anatomy of a frog using tangible AR. In *Proceedings of the Conference on Mensch und Computer*. 303–307.
74. Peck T, Seinfeld S, Aglioti S, et al. (2013) Putting yourself in the skin of a black avatar reduces implicit racial bias. *Consciousness and cognition* 22: 779–787.

75. Qu C, Ling Y, Heynderickx I, et al. (2015) Virtual bystanders in a language lesson: examining the effect of social evaluation, vicarious experience, cognitive consistency and praising on students' beliefs, self-efficacy and anxiety in a virtual reality environment. *PloS one* 10: e0125279.
76. Quintín E, Sanz C, Zangara A (2016) The impact of role-playing games through Second Life on the oral practice of linguistic and discursive sub-competences in English. In *2016 International Conference on Collaboration Technologies and Systems (CTS)*. 148–155.
77. Ratan R, Beyea D, Li B, et al. (2020) Avatar characteristics induce 'users' behavioral conformity with small-to-medium effect sizes: A meta-analysis of the proteus effect. *Media Psychology* 23: 651–675.
78. Redondo B, Cózar-Gutiérrez R, González-Calero JA, et al. (2020) Integration of augmented reality in the teaching of English as a foreign language in early childhood education. *Early Childhood Education Journal* 48: 147–155.
79. Ripka G, Grafe S, Latoschik ME (2020) Preservice Teachers' encounter with Social VR–Exploring Virtual Teaching and Learning Processes in Initial Teacher Education. In *SITE Interactive Conference*. 549–562. Association for the Advancement of Computing in Education (AACE).
80. Sherman WR, Craig AB (2003) *Understanding virtual reality*. San Francisco, CA: Morgan Kauffman.
81. Shih YC, Yang MT (2008) A collaborative virtual environment for situated language learning using VEC3D. *Educational Technology and Society* 11: 56–68.
82. Shih YC (2015) A virtual walk through London: Culture learning through a cultural immersion experience. *Computer Assisted Language Learning* 28: 407–428.
83. Skarbez R, Frederick PB, Mary CW (2018) Immersion and Coherence in a Stressful Virtual Environment. In *Proceedings of the 24th Acm Symposium on Virtual Reality Software and Technology*. 1–11.
84. Slater M, Wilbur S (1997) A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators and Virtual Environments* 6: 603–616.
85. Slater M (1999) Measuring presence: A response to the Witmer and Singer presence questionnaire. *Presence* 8: 560–565.
86. Slater M (2003) A note on presence terminology. *Presence connect* 3: 1–5.
87. Sanchez-Vives MV, Mel S (2005) From presence to consciousness through virtual reality. *Nat Rev Neurosci* 6: 332–339.
88. Slater M (2009) Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philos T R Soc B* 364: 3549–3557.
89. Tulodziecki G, Grafe S (2012) Approaches to learning with media and media literacy education—trends and current situation in Germany. *Journal of Media Literacy Education* 4: 5.
90. Tulodziecki G, Grafe S, Herzig B (2019) *Medienbildung in Schule und Unterricht: Grundlagen und Beispiele*. UTB GmbH.
91. Vate-U-Lan P (2012) An augmented reality 3d pop-up book: the development of a multimedia project for English language teaching. In *2012 IEEE International Conference on Multimedia and Expo*. 890–895.

92. Vedadi S, Abdullah ZB, Cheok AD (2019) The Effects of Multi-Sensory Augmented Reality on Students' Motivation in English Language Learning. *In 2019 IEEE Global Engineering Education Conference (EDUCON)*. 1079–1086.
93. Wang CX, Calandra B, Hibbard ST, et al. (2012) Learning effects of an experimental EFL program in Second Life. *Educational Technology Research and Development* 60: 943–961.
94. Wang YF, Petrina S, Feng F (2017) VILLAGE—V irtual I mmersive L anguage L earning and G aming E nvironment: Immersion and presence. *British Journal of Educational Technology* 48: 431–450.
95. Wienrich C, Johanna G (2020) AppRaiseVR—An Evaluation Framework for Immersive Experiences. *I-Com* 19: 103–121.
96. Wienrich C, Döllinger NI, Hein R (2020) Mind the Gap: A Framework (BehaveFIT) Guiding The Use of Immersive Technologies in Behavior Change Processes. *arXiv preprint arXiv:2012.10912*.
97. Wienrich C, Eisenmann M, Latoschik ME, et al. (2020) CoTeach - Connected Teacher Education. *VRinSight Greenpaper* 53–55.
98. Witmer BG, Singer MJ (1998) Measuring presence in virtual environments: A presence questionnaire. *Presence* 7: 225–240.
99. Wohlgenannt I, Simons A, Stieglitz S (2020) Virtual Reality. *Business and Information Systems Engineering* 5: 455–461.
100. Wu MH (2019) The applications and effects of learning English through augmented reality: A case study of Pokémon go. *Computer Assisted Language Learning*, 1–35.
101. Yang MT, Liao WC (2014) Computer-assisted culture learning in an online augmented reality environment based on free-hand gesture interaction. *IEEE T Learn Technol* 7: 107–117.
102. Yeh HC, Tseng SS, Heng L (2020) Enhancing EFL students' intracultural learning through virtual reality. *Interactive Learning Environments*, 1–10.
103. Zhang X, Zhou M (2019) Interventions to promote 'learners' intercultural competence: A meta-analysis. *International Journal of Intercultural Relations* 71: 31–47.



AIMS Press

©2021 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)