Augmented Reality for Early Language Learning: A Systematic Review of Augmented Reality Application Design, Instructional Strategies, and Evaluation Outcomes

Min Fan[1] , Alissa N. Antle[2], and Jillian L. Warren[2]

Abstract

In this article, we present a systematic review of literature on augmented reality (AR) supported for early language learning. We analyzed a total of 53 papers from 2010 to 2019 using qualitative analysis with complementary descriptive quantitative analysis. Our findings revealed three main AR learning activities: word spelling games, word knowledge activities, and location-based word activities. Our findings also uncovered five main design strategies: three-dimensional multimedia content, hands-on interaction with physical learning materials, gamification, spatial mappings, and location-based features. Several combinations of design and instructional strategies tended to be effective: Learning gains were enhanced by using three-dimensional multimedia with advanced organizers (presentation strategy) and/or using location-based content with learners' self-exploration (discovery strategy); and motivation was enhanced by using game mechanisms with discovery strategy. We suggest that

[1]School of Animation and Digital Arts, Communication University of China
[2]School of Interactive Arts and Technology, Simon Fraser University, Surrey British Columbia, Canada

Corresponding Author:
Min Fan, School of Animation and Digital Arts, Communication University of China, No. 1 Dingfuzhuang East Street, Beijing 100024, China.
Email: mfan@cuc.edu.cn

future designers of AR early language applications should move beyond these basic approaches and consider how unique benefits of AR may be applied to support key activities in early language learning while also considering how to support socio-technical factors such as collaboration between teachers and learners and different learning contexts. We conclude with a discussion of future directions for research in this emerging space.

Keywords
augmented reality, early language learning, systematic review, augmented reality design, instructional strategies, evaluation outcomes

The ability to read and spell is a core competency that can greatly impact individuals' lifelong emotional, educational, and economic outcomes (Earnshaw & Seargeant, 2005). Language learning is a cognitive developmental process and is acquired in stages over time (Ehri, 2014; Frith, 1985). In alphabetic languages, such as English and Spanish, learners need to first learn *phonological awareness* (i.e., the ability to manipulate sound units in oral language) and *the alphabetic principle* (i.e., the rules of how letters are correspondence to sounds; Vellutino et al., 2004). Once learners master these first skills, they can decode words accurately and fluently, and then begin to focus on reading comprehension. However, when learners do not master these core skills, they may show various degrees of difficulties in reading or spelling (Dehaene, 2009). When these difficulties are severe, this is called *dyslexia*. Therefore, the early language skills that are foundational for word decoding are extremely important for learners of alphabetic languages.

To support effective early language learning (i.e., language knowledge acquisition mainly related to letters and word decoding rather than comprehension), many computational systems have been designed. Systems may include graphic user interfaces (GUIs,[1] e.g., Kast et al., 2007; Sysoev et al., 2017) and tangible user interfaces (TUIs,[2] e.g., Goh et al., 2012; Hengeveld et al., 2013). These digital tools are advantageous in supporting literacy learning due to their cost-effectiveness (e.g., available to learners who cannot access to nor afford in-person instructions), multimedia teaching approach, motivational game mechanisms, and support for collaborative learning (Mioduser et al., 2000; Nicolson et al., 2000). Compared to GUIs, TUIs make it easier for learners to position, organize, or trace the physical letters while hearing associated sound changes along with other informational cues (Hengeveld et al., 2013), which can leverage multiple senses (particular tactile and kinesthetic senses) in supporting language learning.

New technologies including mobile tablets and augmented reality (AR) headsets have enabled researchers to turn their attention to using AR technology for language learning. AR is a three-dimensional (3D) technology that enhances the user's sensory perception of the real world by overlaying it with a contextual layer of digital information (Azuma, 1997). Compared to other types of user interfaces such as GUIs and TUIs, the unique affordances of AR applications lie in they can (a) directly augment the reality by providing a digital overlay on top of (rather than tangible) learning materials, which may better draw learners' attention to relevant visual-audio information (Fan et al., 2018; Radu, 2014); (b) associate learning with specific contexts which enables the situated learning opportunity (Hsu, 2017; Santos et al., 2016); (c) hands-on interaction with physical learning materials as well as seeing associated digital feedback on top of hands-on actions may leverage learners' embodied knowledge in understanding abstract concepts (Barreira et al., 2012), and (d) run in low-cost handheld devices that can be easily deployed at schools or home. Therefore, with the commercial availability of portable devices, AR technology that combines the advantages of GUIs and TUIs may have a great potential to support learning.

Many studies have presented the design and evaluation of specific AR applications that support language learning, but we have not seen any systematic reviews on AR applications for early language learning. There are some reviews that have synthesized the results of previous studies regarding educational AR to identify broader research trends, that is, AR technology in educational settings (Akçayır & Akçayır, 2017; Bacca et al., 2014; Radu, 2014), and several have covered computational technology for language learning (Hung et al., 2018; J. Yang, 2013; X. Yang et al., 2018). However, we have not seen any reviews which have specifically discussed learning activities, design strategies, and instructional strategies that are frequently used in AR applications for early language learning and how these factors may be correlated or causally related to positive learning process and outcomes. Since early language learning plays a vital role in language development and may affect the achievement of many other essential abilities, research on AR design for early language learning may contribute to design knowledge on how to leverage the use of this potential technology to support learning in this field.

The research presented in this article addresses the gap for understanding the potential intertwined factors of AR application design that may support learners' early language acquisition. Through a systematic review, we aimed to address the research gap by seeking answers to four main research questions: (RQ1) What are the main kinds of activities in AR applications for early language learning? (RQ2) What are the main design strategies in the AR applications? (RQ3) What are the main instructional strategies in the AR applications? (RQ4) What are the main evaluation goals, methods and outcomes of the AR applications?

Through the systematic review, we provide designers, researchers, and practitioners with an overview of the current landscape of AR for early language acquisition. We also aim to provide practical and forward-looking knowledge by (a) looking at areas of early language learning where combinations of design strategies and instructional strategies were shown through evaluations to be effective as positive exemplars; (b) providing design strategies that could be used to create new forms of AR applications to support a variety of early language learning activities; and (c) providing future research directions on exploring the promising design and instructional strategies that may support language learning with AR applications. Overall, this systematic review contributes to the design, instructional, and research knowledge of AR applications designed for early language learning for alphabetic language learners.

Related Work

Types of AR

To overlay digital information onto physical objects, AR technologies must track objects in the physical world in real time. As such AR interfaces are often described based on the specific techniques used for tracking objects in the physical world. Several researchers classified AR interfaces into two levels based on their tracking techniques: (a) image-based AR including marker-based and markerless solutions that use image recognition techniques to track an object and its position and (b) location-based AR that uses position data to identify an object and its position (Cheng & Tsai, 2013; Koutromanos et al., 2015). AR applications can run on a computer-webcam system (Hornecker & Dünser, 2007), head-mounted display (Milgram et al., 1995), handheld display (Juan et al., 2010), or projector-camera system (Silva et al., 2013). We focused on all types of the AR interfaces in this review.

Reviews on Computer-Assisted Language Learning

Several narrative reviews[3] have been performed that focused on computational tools for literacy acquisition. Vernadakis et al. (2005) presented a narrative review to discuss the potential benefits of computer-assisted instruction on cognitive, emotional, linguistic, and literacy skills of preschool children in the classroom. This review indicated that GUI-based instruction could improve preschool children's phonological awareness, word recognition, writing ability by increasing the children's attention span with animated and interactive multimedia contents, and enabling children's self-directed learning that best matched their own learning pace. The authors also pointed out that the role for teachers in the learning process is often overlooked. J. Yang (2013) reported a narrative review that focused on mobile-based pedagogical applications for

language learning. The author categorized five main use scenarios with mobile technology in language learning, including Short Message Service, Microblogging, Ambient Intelligence and Augmented Reality, GPS, and Tablet Computing. Yang illustrated that the greatest advantage of mobile-based AR is its ability to provide digital overlay on the real world and support context-aware language learning.

Several reviews used a descriptive[4] or statistical approach[5] to systematically search, filter, and classify the empirical studies related to the use of computer-assisted tools in supporting language learning. Colwell and Hutchison (2015) presented a systematic review of 11 empirical studies on digital tools used in K-5 language arts classrooms published between 2000 and 2013. The review suggested a variety of digital tools (e.g., blog, iPad apps, online games) could be used together in the classroom to support children's reading comprehension, discussion and collaborative learning. In addition to the general benefits (i.e., motivation, interaction, multimedia contents, collaboration) offered by the digital tools, the authors also advocated to integrate digital technology into teachers' daily instruction. X. Yang et al. (2018) focused on computer-based reading instructions for K-12 learners and the associated underlying reading theories. This article examined 70 articles published between 2004 and 2015 in two journals on reading and literacy learning. The results revealed the primary learning focus with technology were vocabulary and reading comprehension acquisition, and the major functions of technology in language acquisition were to increase learners' reading motivation, present them with multi-modal information, and promote their collaborative learning, based on the theories of Reading Motivation, Dual Coding, and Social Constructionism, respectively. Hung et al. (2018) presented a systematic review on empirical evidence on the use and impacts of digital games in language education. A total of 50 studies from 2007 to 2016 were systematically analyzed. The results suggested that digital games were often used in applications designed for second language learning, and *immersive games* (including characters and narratives) and *tutorial games* (learning through drills and practice) were the two most popular games genres in language learning domain. Zucker et al. (2009) conducted a meta-review to investigate the effects of ebooks on language development for elementary learners. This review included 23 studies published in the databases of PsycINFO, ERIC, and Web of Science between 1997 and 2007. The result suggested the potential benefits of ebooks' hypermedia features (e.g., animations) in supporting young children's short-term comprehension skills.

In summary, the aforementioned reviews revealed common strategies (i.e., multimedia contents, interaction, games, collaboration) of computational tools used for language acquisition and education along with their potential benefits (i.e., increasing attention, promoting motivation, and enabling learner-centered/ collaborative learning) in supporting a variety of language learning activities (i.e., text comprehension, vocabulary, phonological awareness/speaking, listening reading, writing). A few reviews also discussed the challenges of designing

and employing these computational tools in the context (i.e., considering the learning curve and specific requirements of teachers). However, the learning scenarios analyzed in these reviews were not specifically focused on early language learning stage or AR applications.

Reviews on AR Learning Technology

Several narrative reviews have emphasized AR technology for learning (not specific for language learning), and have primarily focused on the advantages, challenges, and effectiveness of using AR technology in general educational contexts. Wu et al. (2013) provided a comprehensive narrative review on AR's affordances, instructional strategies, challenges, and future research directions. The authors summarized the main affordances of AR lie in its 3D visualizations which enable authentic, situated, collaborative, and immersive learning opportunities. They also brought up the issues of AR's technical failure, the lack of instructional design, the inflexibility of the content updates, and the cognitive overload caused by the amount of information and the complexity of tasks the learners encountered. Radu (2014) presented a narrative review of 26 empirical studies comparing AR to non-AR applications for education. The author identified the positive impacts of educational AR technology on learning motivation, cognitive development (e.g., spatial ability, memory, attention), and learning strategy (e.g., collaboration) as well as negative impacts such as extra cognitive load, ineffective classroom integration, learner differences, and usability issues. The author constructed a heuristic questionnaire that provides the justifications for the use of AR applications in educational contexts for designers. Several considerations included whether the application makes concepts easier to understand, presents relevant educational information at the appropriate time and place, directs learner's attention to important aspect, or enables learners to physically interact with educational materials.

Several systematic and meta-reviews were also conducted to the most frequently mentioned advantages and challenges of AR applications in supporting education and discuss the design implications for future AR applications. Bacca et al. (2014) synthesized research from 32 studies published from 2003 to 2013 in six journals in Educational Technology with high impact factors. The results showed that the top reported advantages of AR in educational settings included learning gains, motivation, interaction, collaboration and low costs. Commonly mentioned limitations included unstable tracking, too much attention on visual information, and inflexibility of the content updates. The authors also suggested future research to explore AR applications designed for personalized learning styles including learners with special needs as well as AR's impact on learning effectiveness with larger population sizes. Akçayır and Akçayır (2017) presented a systematic review of 68 research articles on educational AR applications for K-12 learners published in the SSCI-indexed journals before 2016. The authors

coded the advantages and challenges of using AR in educational settings and most of the results were consistent with those found in Bacca et al.'s study. The results showed that AR technology could promote learning outcomes and increase learning motivation by leveraging the use of well-integrated and relevant augmented visualizations that served as the learning guidance. Some noted challenges imposed by AR were usability issues such as unstable tracking quality. Koutromanos et al. (2015) reported a systematic review that focused on handheld AR games in the context of formal and informal educational environments. The authors analyzed seven articles between 2000 and 2014 in ScienceDirect and ERIC. The authors argued that games could encourage collaboration between learners, serve as a scaffolding for reading, and provide high levels of engagement. Li et al. (2017) focused on AR games for learning and presented a literature review of 26 papers published from 2010 to 2016. The authors identified the most commonly used game elements were *quizzes* and *goal-setting*, while the frequently used AR features were *extra instructional materials* and *3D models*. The authors also provided five suggestions for designing AR games, including involving learners in the design process, using clearer learning objectives, identifying effects of AR features, studying the game mechanisms, and encouraging social interaction. Santos et al. (2014) investigated the design and evaluation of educational AR applications for K-12 learners. They conducted a meta-analysis review on 87 articles published in the IEEE Xplore Digital Library before 2012. The meta-analysis results showed the affordances of AR included *real-world annotation*, *contextual visualization*, and *visual-haptic visualization*.

To sum up, the aforementioned reviews discussed the general affordances and challenges of using AR in educational contexts. In addition to the common features and benefits shared with other computational approaches, the unique beneficial features of AR applications for learning in general seem to include 3D annotations and visualizations on the real world, visual-haptic interactions with physical learning materials, and context-aware (location-based) learning materials. Several common issues such as unstable tracking quality and inflexibility of content updates were also discussed. However, no reviews specifically examined the use of AR technology in supporting language learning, particularly in the foundational stage required for word decoding that is crucial for all types of language learners. We address this gap with this review research.

Research Methodology

Our objective was to examine the ways in which AR is an emerging research area that may provide support for early language learning and provide a snapshot to guide future design and research. Given the nascence of this research area, we did not and could not aim at examining any variables, correlations, or

theories (meta-analysis). Therefore, a systematic review approach is most appropriate for the current stage of our research (H. Yang & Tate, 2009).

Paper Quality

Since the use of AR for language learning is a newly emerging field, and one of our research goals was to inform the design of AR language learning applications, all within-scope peer-reviewed research papers, including demonstration and extended abstract papers, were included in this review.

Manuscript Searching and Filtering Process

In this review article, we included English language peer-reviewed articles that were published between January 2010 and 2019. We decided to use online database searches as a primary literature collecting approach because this research area is emergent and its publication channels are still scattered. Seven well-known online research databases were used to find relevant literature sources based on previous research (Ibáñez & Delgado-Kloos, 2018). These databases are related to education and technology, including ACM Digital Library, ERIC, PsycINFO, IEEExplore, Web of Science Core Collection, ScienceDirect, and Springer Link.

Based on our research goal, two sets of keywords were developed: (a) technology-related keywords, including *augmented reality*, and *AR*; as well as (b) language-related keywords, including *language*, *book*, *literacy*, *read*, *spell*, *speak*, *write*, *letter*, *alphabet*, *alphabetic*, *vocabulary*, *word*, and *English learning* following from (Hung et al., 2018). These two sets of search terms were run in combination with Boolean operators, with the AND operator used between the sets and the OR operator used within the set. We conducted keyword and abstract searches across all seven databases. The last search was conducted on December 30, 2019. The search produced 1,247+ results from the aforementioned search terms and designated time period, including 302+ duplicates, which were deleted. This left us with 945 articles.

In the *second round* of the filtering process, we reviewed *the title and/or abstract* of all 945 papers based on the criteria listed in Table 1, in order to quickly filter and remove the unrelated ones. The title and/or abstract of the 945 articles were examined by the first author to determine whether they were suitable for the purposes of the study. One co-author independently reviewed approximately 20% of the articles (a total of 189 papers including 27 papers randomly selected from each database) to confirm the reliability of the selection. The intercode agreement rate for coding was 95.2%. Disagreements between the two coders were resolved through discussion and further review of the disputed studies. There were 138 papers left after the second-round selection.

Table 1. Inclusion and Exclusion Criteria.

Inclusion criteria	Exclusion criteria
(a) Peer-reviewed article.	(a) Review or meta-review articles.
(b) Available in full-text in English.	(b) Studies that mentioned AR and learning, but were about learning other topics such as science or history rather than early language learning.
(c) Described AR applications supporting early alphabetic language learning as a primary component. Languages included English and similar *alphabetic languages* (Apel, 2011).	(c) Described AR applications supporting deaf learners learning sign languages.

Note. AR = augmented reality.

In the *third round*, all remaining articles were *read in their entirety* in order to verify that these articles met the criteria in Table 1. The second author independently reviewed approximately 20% of the 138 articles. The intercode agreement was initially 93.5%, and this was brought to 100% after discussion. A total of 53 articles met the criteria for inclusion in the final review.

To collect as many papers as possible, we also across checked the references of previous systematic reviews and added seven articles that met our criteria. Therefore, there were a total of 53 papers included in our final review.

Developing Codes/Scheme

To code the studies in a systematic way, a coding scheme was developed (Tables 2 to 5). The coding scheme consisted of four clusters of categories (i.e., *AR activities*, *AR design strategies*, *AR instructional strategies*, and *evaluation goals, methods and outcomes*), with each including several main and subcategories. The initial coding scheme was developed in accordance with previous theories and research on AR design and learning (Akçayır & Akçayır, 2017; Ibáñez & Delgado-Kloos, 2018; Wu et al., 2013; X. Yang et al., 2018) and the research questions of this study. During the systematic review process, some subcodes emerged and the coding schemes were refined in order to reflect the emerging information.

The first coding category (*AR activities*) was to answer the RQ1 (*What are the main kinds of activities in AR applications for early language learning?*). We first wanted to know whether there were common learning activities with these AR applications. The second coding category (*design strategies*) focused on RQ2 (*What are the main design strategies in the AR applications?*). We aimed to understand if these AR applications shared similar design strategies. The third coding category (*instructional strategies*) focused on RQ3 (*What are the main instructional strategies in the AR applications?*). We focused on the instructional strategies and

Table 2. Coding Scheme for RQ1: Types of AR Learning Activities.

Preset categories	Operational definitions (*emergent codes are in italics*)	Examples
RQ1: Types of Learning Activities	Types of AR learning activities refer to the categorization of AR learning activities based on their learning focus and interactional process. *1st code (augmented word spelling games)—AR applications have hands-on interaction with physical objects (e.g., letters, cards, cubes) which are augmented through displaying digital contents on a screen, and include game activity that focuses on spelling.*	• AR Magic English (Pu & Zhong, 2018) • AR PhonoBlocks (Fan et al., 2018) • AR game (Juan et al., 2010)
	2nd code (paper-based word knowledge visualization activity)—RQ1-AR applications have hands-on interaction with physical objects (e.g., cards, books) which are augmented through displaying digital contents on a screen, and may include any non-spelling activities.	• AR 3D Learning Objects (Luna et al., 2018) • AR 3D Pop-up Book (Juan et al., 2010) • MOW (Barreira et al., 2012)
	3rd code (location-based word knowledge visualization activity and/or spelling games)—AR applications that either do not contain physical objects but just using mobile devices (phones or tablets) to augment the real world or contain 3D everyday objects located in the real world.	• Vocabulary AR game (Hsu, 2017) • Situated Vocabulary Learning System (Santos et al., 2016) • ARbisPictus (A. Ibrahim et al., 2018)

relevant background information involving instructional goals, instructional materials, and if the instruction was offered by teachers or not because these factors may influence the instructional strategy. The fourth coding category (*evaluation goals, methods and outcomes*) emphasized RQ4 (*What are the main evaluation goals, methods and/or outcomes of the AR applications?*). We investigated the evaluation goals, methods, and outcomes of the AR applications. We also examined the background information in terms of the evaluation scale (e.g., participant sample size) and context (e.g., study place such as school or lab), which may impact the choices of specific methods used in evaluations.

The analysis of the four clusters of coding themes not only allowed us to look for answers to each individual RQ but also enabled us to (a) look at areas of

early language learning where combinations of design strategies and instructional strategies were shown through evaluations to be effective as positive exemplars, (b) look for potential design strategies and create new forms of AR applications that may be used in the school context to support a variety of early language learning activities, and (c) look at future research directions on the evaluation of promising design and instructional strategies within AR applications.

The selected articles were analyzed mainly using a qualitative content analysis method, which is a thematic analysis concentrating on the relationship between the articles' content and the shared characteristics on certain themes (Elo & Kyngäs, 2008; Radu, 2014). We also provided complementary descriptive statistical analyses to identify the general trend (i.e., percentage) of the application of the strategies. Two of the authors of this article manually coded the studies separately according to the preset (and evolving) categories and codes.

Our specific coding process followed these seven steps: (a) scanned all the articles to get the general meaning of the whole; (b) selected each article to read and look for the most descriptive corpus sentences related to our preset categories/codes (e.g., we marked the sentences or keywords or rewrote the summary based on the description); (c) coded the text; (d) reread the article and listed the interesting content related to our topic but not in the preset codes, and then added them into the categories; (e) checked for discrepancy, and resolved them through discussion; (f) repeated the aforementioned process for coding the next five articles, repeated until the end; and (g) if any new codes emerged, the corresponding contents in the prior coded articles would be recoded and updated.

Results

Overview of Reviewed Papers

There were a total of 53 papers reviewed from 2010 through 2019 (Online Appendix A). Approximately two thirds ($n = 36$) were published after 2015 and half ($n = 26$) were from IEEExplore database. Around 75% of the articles ($n = 39$) focused on research on mobile-based AR systems while 53% the articles ($n = 28$) defined their designs as games. The most preferred development tools were *Unity* and *Vuforia* plugin ($n = 16$). The learning languages in the studies were mostly English ($n = 45$). The learners in the studies were mostly primary school children ($n = 23$), followed by preschool children ($n = 17$). In the 45 English-focused articles, the 40 articles targeted learners learn English as a foreign language (EFL). Four articles discussed learners with special needs (two for Autism, one for dyslexia, and one for attention deficit hyperactivity disorder). More than half of the articles ($n = 28$) focused on studies to evaluate the learning effects of AR applications, affective outcomes, and/or usability testing; several

($n = 11$, 21%) focused on usability testing results; a few ($n = 6$, 11%) emphasized the investigation of affective outcomes; the rest ($n = 8$, 15%) presented AR prototype design without any studies for evaluation.

RQ1: Types of AR Activities

This section addresses our first research question. We identified 49 applications from the articles reviewed (since a few articles discussed the same AR application). The main activities in the AR applications could be broadly categorized into three main types (Table 2).

The first type focused on *augmented word spelling games* ($n = 9$, Figure 1, *left*). In this type, learners can (a) use physical 2D lettercards or 3D letters to construct a word, phrase or sentence, and view the associated digital letters, sounds, and models/animations of the word on the screen (Boonbrahm et al., 2015; Fan et al., 2018; Zhenming et al., 2017); or (b) place a 2D picture flashcard under a tablet to trigger several 2D digital letters on the display, and then reorder the digital letters through the direct dragging on the screen (Y. Chen et al., 2017; Pu & Zhong, 2018). For example, in the AR application presented by Boonbrahm et al. (2015), learners can mix letters to make a word; and if the word matches the name of the animal, that 3D animal will pop up and start doing some activity. In another application (Pu & Zhong, 2018), learners can place a picture card (e.g., a *plane* card) under the tablet, and the letters of the word (e.g., *p, l, a, n, e*) are randomly shown on the screen; learners need to drag the digital letters into the corresponding letter box to complete the word spelling.

The second type, *paper-based word knowledge visualization activity* ($n = 30$, Figure 1, *middle*), emphasizes the visualization of extended language knowledge on 2D physical flashcards or books. This kind of application usually augments physical flashcards or books by displaying rich digital content such as 3D animations, video, 2D/3D letters, and/or audio on a display (Cieza & Lujan, 2018; Luna et al., 2018; Rambli et al., 2013). This multimedia content is used to explain the spelling, sounds, meanings, or other reverent knowledge of letters and words. For example, in the *Word Worth Learning* application (Luna et al.,

Augmented word spelling games | Paper-based word knowldge visualization activity | Location-based word knowldge visualization activity and/or spelling games
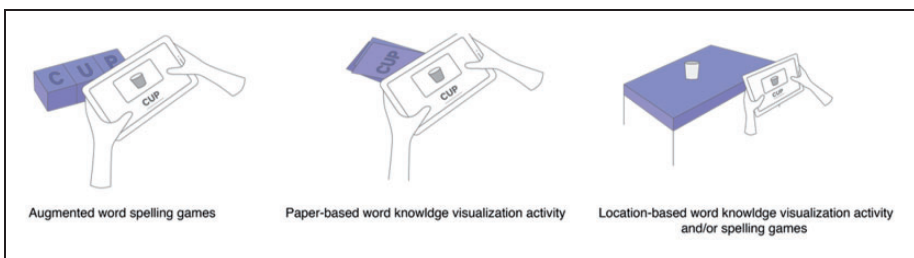
Figure 1. Three Main Types of AR Learning Activities.

2018), when learners place a picture card under a tablet, the associated 3D letters, sound, and 3D animation are shown on top of the card. A few applications of this kind also included several extended activities. Different to the first type focused on spelling, these activities were mainly about *word-meaning association* or *word comprehension in text*. In the AR applications in (Barreira et al., 2012; Chang et al., 2011), learners needed to correctly match a picture card with a word card; if the matching is correct, the augmented 3D object and 2D rewarding feedback text are shown on the screen. In the AR pop-up books, learners could press the augmented hotspots (e.g., "!" symbol) to listen to the animal sounds (Mahadzir & Phung, 2013) or the dialogues from the narrator (Vate-U-Lan, 2011).

The third type, *location-based word knowledge visualization activity and/or spelling games* ($n = 10$, Figure 1, *right*), contains similar activities with the previous two types. However, the activities in this kind are location-relevant. That is, augmented content is triggered by specific large-scale real-world scenes (e.g., desk surfaces or buildings; Ho et al., 2017; Lee et al., 2017), or small-scaled everyday objects placed in the environment (e.g., garbage can, table, cup; Hsu, 2017; A. Ibrahim et al., 2018; Santos et al., 2016). For example, in the *UL-IAR* application, learners received a set of real-time English quizzes in real-life contexts based on the GPS positioning function (Ho et al., 2017). In the *ARbis Pictus*, learners saw the augmented words on top of the 3D objects through a head-mounted display (A. Ibrahim et al., 2018).

RQ2: Design Strategies

We identified five main design strategies that were frequently used in AR design for early language learning (Table 3). They are related to: (a) 3D multimedia for letters, sounds, and meanings ($n = 49$, 100%); (b) hands-on interaction with physical letters, cards, or objects ($n = 44$, 90%); (c) gamification involving word building games ($n = 27$, 55%); (d) congruent spatial mappings between physical representations (usually inputs) and digital representations (usually outputs; $n = 21$, 43%); and (e) location-based features ($n = 9$, 18%).

1. *3D multimedia content*: 3D digital representations such as 3D letters, word models, or animations with sounds were frequently used to visualize vocabularies with concrete (rather than abstract) meanings. We found almost half of the applications contained 3D models and one-third used 3D animations. The 3D models and animations together with sounds were mostly used to demonstrate word spelling, sounds, and meanings. Learners could either inspect the 3D virtual objects/letters of the word (e.g., *plane → a 3D virtual plane*) from a variety of different perspectives (Pu & Zhong, 2018; Rambli et al., 2013) or watch a sequence of looped 3D animations to understand the

Table 3. Coding Scheme for RQ2: Design Strategies and Features.

Preset categories	Preset and emergent codes (emergent codes are in italics)	Operational definitions (Emergent definitions are Italic)	Examples	Based on theories or previous research
RQ2: Design strategies	Design strategies	Design strategies refer to the high-level and specific strategies in digital representations (i.e., multimedia content, including identifying visual—auditory modalities and 2D/3D, pictures/animations/videos), interactions (i.e., gamification, including narrative and non-narrative game genres; Hung et al., 2018), physical representations (i.e., physical learning materials), and other features related to language learning activities (e.g., mappings between physical and digital representations, location-based features; Fan et al., 2017).	**Multimedia content:** using 3D model that can be viewed from 360° to demonstrate word meaning (Luna et al., 2018) **Hands-on interaction with physical learning materials:** 3D letters (Fan et al, 2018); 3D every-day objects (Santos et al., 2016) **Gamification:** word spelling games (Pu & Zhong, 2018); word-object matching games (Barreira et al., 2012) **Congruent spatial mappings between physical and digital representations:** spelling games with logo-based notches (Fan et al, 2018a) **Location-based features:** linking indoor real-world scenes to word spelling games (Lee et al., 2017) Dynamic color cues: AR PhonoBlocks (Fan et al, 2018)	Dual Coding Theory (Sadoski & Paivio, 2013) Cognitive Theory of Multimedia learning (Mayer, 2014) Situated Cognition (Brown et al., 1989) Cognitive Load Theory (Sweller, 1988) Embodied Cognition (Antle, 2013)
	Other design aspects	*Other design aspects include the emerging design facts or issues identified in the review.* *• Uppercase or lowercase letters* *• Color cues*		

use of the words in contexts (e.g., *cat → a running cat*; *hit → hit a ball*; Boonbrahm et al., 2015; Luna et al., 2018).

2. *Hands-on interaction with physical learning materials*: Almost all the AR applications consisted of multiple hand-sized and light physical learning materials, and most of them were 2D flashcards. A large proposition of the applications (*n*=37) used 2D flashcards such as letters, pictures (He et al., 2014; Li et al., 2015), or textbook pages (Kucuk et al., 2014; Rambli et al., 2013) while several applications (*n*=9) used 3D generic cubes, letters, or everyday objects such as pencils and backpacks (Hsu, 2017; Sansosti et al., 2004).

3. *Gamification*: More than half of the AR applications utilized gamification. The most common games were *spelling games* (e.g., a set of quizzes for word spelling tasks; Pu & Zhong, 2018), followed by *matching games* (e.g., matching a picture with a word; Barreira et al., 2012), *collecting games* (e.g., finding out the target object associated with a word in real-life situations, usually in location-based AR applications; Hsu, 2017), and *puzzle games* (e.g., placing the events in chronological order correctly, usually in pop-up AR books; Vate-U-Lan, 2011).

4. *Congruent spatial mappings between physical and digital representations*: Around 40% of AR applications emphasized the spatial characteristics of AR, with many leveraging physical affordances of learning materials or digital scaffoldings in the interface design in the spelling games, and a few utilizing physical space to teach learners preposition words (e.g., on/under, left/right). For example, in five applications that contained word spelling games, letterbox or underlines were intentionally designed in the interfaces, which gave learners an explicit clue about the number of the letters in the spelling task (e.g., f i s h, for *fish*). One application specifically used logo-based notches as physical affordances so that children were only enable to connect words in a certain order to make a sentence (Fan et al, 2018a). Three AR applications specifically focused on teaching children preposition words (Dalim et al., 2016; Hsieh et al., 2014; Vate-U-Lan, 2011).

5. *Location-based features*: Approximately 20% of AR applications utilized location-relevant learning contents. Seven applications contained indoor activities (Y. Chen et al., 2017; Hsu, 2017; Lee et al., 2017), while two had an outdoor activity (Ho et al., 2017; Wu, 2019). Three applications directly augmented real-world scenes (Ho et al., 2017; Lee et al., 2017), while the other six augmented everyday objects (A. Ibrahim et al., 2018; Santos et al., 2016).

We also noticed other design decisions. First, we found the various uses of uppercase and lowercase letters among digital content and physical representations in AR applications. Nine applications included uppercase letters, eight applications used lowercase letters, and five applications contained both

uppercase and lowercase letters. Second, we found that most applications contained rich colors to attract learners' interests. Two designs specifically used augmented color cues to promote children's learning of the alphabetic principle (Fan et al., 2018) and sentence constructions (Fan et al, 2018a). Colors were used to attract children's attention to the stable patterns of words or certain structures of sentences.

RQ3: AR Instructional Strategies

Context for Instructional Strategies. Since we focus on early language learning, all the articles reviewed emphasized letters or word levels but with various focuses. Around two thirds of the articles ($n = 33$, 62%) discussed vocabulary learning with the central focus on spelling, sounds and meanings using the *whole word* approach (A. Ibrahim et al., 2018; Luna et al., 2018). Several articles ($n = 10$, 19%) involved word learning and practical use in sentences and texts. They also used the whole word approach, but focused on learning words in sentences and paragraphs (Dalim et al., 2016; Solak & Cakir, 2015). A few articles ($n = 5$, 9%) dealt with single alphabet reading and writing (Bhadra et al., 2016; Rambli et al., 2013). The rest articles ($n = 5$, 9%) targeted on phonological rules, that is, the alphabetic principle (Fan et al., 2018); vowel sounds (Cieza & Lujan, 2018); phonic training (Limsukhawat et al., 2016); onset and rime training (Juan et al., 2010); and prefix, root, and suffix (Wu, 2019). In addition to the five articles specifically focused on phonological training, six articles also integrated some phonological trainings into their learning activities as a secondary point of training. For example, in the AR pop-up book, students learned the words *seed* and *seat* which share the same onset and syllables when reading the story (Vate-U-Lan, 2011); in the alphabet training, learners could also learn the associated letter sound (Rambli et al., 2013). The majority of the studies emphasized reading and spelling while three focused on handwriting (Ati et al., 2018; Bhadra et al., 2016; Boonbrahm et al., 2015).

In more than half of the articles ($n = 30$, 57%), the use of the AR applications was completed under the guidance of teachers or parents. The majority of the instructions ($n = 37$, 70%) were individual-based, and several were designed for small groups ($n = 8$, 15%) and for classrooms ($n = 8$, 15%). The word themes used most frequently within the AR applications were those contained concrete meanings, including *animals* (e.g., *cat*, *fish*), *fruits* (e.g., *apple*, *banana*), *common objects* (e.g., *table*, *door*), *transportation tools* (e.g., *car*, *plane*), *colors* (e.g., *red*, *green*), and *body names* (e.g., *hands*, *eyes*). A few applications also included spatial words such as *proposition words* (e.g., *under*, *in*).

AR Instructional Strategies. Based on previous research (Ibáñez & Delgado-Kloos, 2018), the studies reviewed could be grouped into three categories: *instruction through presentation* (i.e., teacher-centered informal instruction), *instruction*

through discovery (i.e., learner-centered comprehensive instruction), and *collaborative learning* (i.e., learner-centered group studies; Table 4). We categorized the articles mostly based on the descriptions of the practical use of the AR application (e.g., study procedure, application design, use scenario), which reflected both the built-in design features of the application as well as the intended use of the application in contexts.

The *instruction through presentation* strategy is based on Ausubel's Meaningful Learning Theory, and it emphasizes that learning a subject should be presented first by advanced organizers (usually teachers) and then progressively differentiated in terms of details (Akdeniz, 2016). Presentation instructional strategy was used in 15 articles. More than half of the instructions ($n = 9$) were conducted in groups. A typical instructional process was that a teacher led the instruction and used the AR application for demonstration, and then learners had a chance to use the system to practice the learnt knowledge under the guidance of the teacher (Barreira et al., 2012; Silva et al., 2013; Solak & Cakir, 2015).

The *instruction through discovery* strategy is based on Bruner's Theory of Development which assumes that a learner constructs his or her knowledge through discovering for themselves as opposed to being told (Akdeniz, 2016). Discovery instructional strategy was used in 35 articles. In these instructions, learners were given an AR application to conduct self-directed learning through a set of games or application guided reading activities. Learners received corresponding feedback/hints from the applications. Teachers/parents worked as the activity facilitator to ensure the activity went well (Hsu, 2017; Juan et al., 2010; Pu & Zhong, 2018; Rambli et al., 2013).

Collaborative learning is a form of small group instruction where learners work in a social setting to solve problems (Slavin, 1991). Collaborative learning was used in 15 articles, in which 13 also used discovery instructional strategy. In these learning activities, learners were often assigned tasks (e.g., discussing to answer the questions in pop-up books) and peer discussions were encouraged during the learning process (C. Chen & Wang, 2015; Vate-U-Lan, 2012; Zainuddin & Idrus, 2016).

RQ4: Evaluation Goals, Methods, and Outcomes

Background for Evaluation. Among the 53 articles reviewed, 45 articles conducted evaluation studies. One third of the studies ($n = 16$, 30%) included 10 to 50 participants; followed by those ($n = 9$, 20%) including 50 to 100 participants. The sample sizes in seven studies were above 100 (with the most including 484 participants), and three studies had fewer than 10 (with the fewest including five participants). In the 70% of the articles ($n = 31$) that targeted children, two thirds specifically focused on Grade 1 and Grade 3 learners (aged 6–9 years). Approximate half of the AR instructions ($n = 22$) were conducted with individual participants, nearly 25% ($n = 11$) of the studies were conducted with small

Table 4. Coding Scheme for RQ3: Instructional Strategies.

Preset categories	Preset and emergent codes	Operational definitions (emergent codes are in italics)	Examples	Based on theories or previous research
RQ3: AR instructional strategies	Contexts for instructional strategies	• **Learning goals** mean what specific cognitive achievement the instruction aims to achieve (Akdeniz, 2016; Ehri, 2014). Codes include: ○ Letters/Alphabets ○ Whole word knowledge ○ Word and use in sentences and texts ○ *Phonological knowledge* • *Learning materials include letters, words, sentences used in instructions.*	• **Learning goals:** alphabet reading (Rambli et al., 2013); whole word learning (Luna et al., 2018); whole words in sentences (Solak & Cakir, 2015); phonological knowledge (Limsukhawat et al., 2016) • **Learning contents:** animal words (*cat, fish*)	Connectionist Model (Ehri, 2014)
	Instructional strategies	**Instructional strategies** are instructional methods developed based on learning theories which help teachers to achieve the fundamental aims of instruction (Akdeniz, 2016). Codes include: • Expository presentation (i.e., mainly through teacher's explanation) • Discovery (i.e., mainly through child's free play) • Collaboration (i.e., mainly through children's collaborative learning)	• Expository presentation (Solak & Cakir, 2015) • Discovery (Hsu, 2017) • Collaboration (C. Chen & Wang, 2015)	Ausubel's Meaningful Learning Theory (Akdeniz, 2016) Bruner' Theory of development (Akdeniz, 2016)

groups, and a few ($n = 7$) were conducted with a class group. The majority of the studies ($n = 32$) were conducted in the school context.

Evaluation Goals and Methods. Three main research purposes emerged from the studies reviewed (Table 5). First, the majority of the studies ($n = 28$) focused on investigating the effectiveness of AR applications (mainly learning gains). Second, a few studies ($n = 6$) targeted on exploring the affective aspects (mainly motivation). Third, several studies ($n = 11$) emphasized on the design and system usability evaluation.

More than half of the 45 studies with formal evaluations were quantitative studies ($n = 21$), followed by mixed-method studies ($n = 20$) and qualitative studies ($n = 3$). Four studies only discussed system testing methods such as *heuristics* and *expert review*. Quantitative studies were divided into quasi-experimental studies and experimental design studies. The most common experimental design ($n = 20$) was to compare learners' learning achievements between the AR-based instruction and other comparative instructions (e.g., traditional instruction/a similar AR approach). In these studies, learners were usually assigned to either the AR-based instruction group or the traditional instruction group (between-subject design); Pretests and posttests on certain language-related tasks were administrated before and after the instruction to evaluate learners' learning gains while the five-point Likert-type scale questionnaire mostly based on Keller's ARCS Motivation Model (Song & Keller, 2001) was usually employed after AR usage to investigate learners' motivation and/or perceived usefulness of the instructional methods. In seven studies, there were no matched control groups during the instructional process.

One third of the studies reviewed ($n = 20$) combined quantitative and qualitative research methods for a better understanding of the system effectiveness in supporting language acquisition. These studies usually used an experimental design to evaluate learners' learning gains, while qualitative methods such as observations or interviews were also employed to provide complimentary data with the explanation of the quantitative results.

Three studies employed a qualitative research approach. One exploratory pilot study used interviews to assess the influence of an AR-based English learning app on undergraduate learners' vocabulary learning motivation (Li et al., 2015). Two pilot studies used observations to understand children's general use and learning experience with the system (Dalim et al., 2016; Tang et al., 2019).

Evaluation Outcomes. Thirty-two studies examined learners' language achievements after receiving AR instruction. A majority of the studies reported a positive effect of AR instruction. The positive effects were based on the comparison of learning effectiveness between AR instruction and the traditional instruction, a similar AR approach ($n = 20$), learners' pretest and posttest scores ($n = 10$), or learners' perceived usefulness ($n = 10$). Only three studies reported similar

Table 5. Coding Scheme for RQ4: Evaluation Methods and Outcomes.

Preset categories	Preset and emergent codes	Operational definitions (emergent codes are in italics)	Examples
RQ4: Evaluation goals, methods and outcomes	Background for evaluation	**Learner groups and evaluation context:** mean the learning group sample size and evaluation context. Codes include: • Learner groups: sample size • Evaluation context: classroom, in small groups, individually, school or others	• Learner group: 484 participants (Vate-U-Lan, 2012) • AR classroom-based instruction (Solak & Cakir, 2015) • AR instruction in small groups (C. Chen & Wang, 2015) • AR individual instruction (Pu & Zhong, 2018)
	Evaluation goals and methods	Evaluation goals mean what the main aim and focus of the study. Codes include: • Learning effectiveness: mainly to understand what were the learning efficacy or effectiveness of an AR application, including evaluating learners' learning gains or their *perceived usefulness* of an application. • **Affective outcomes:** mainly to understand what were children's **attitude, learning interests,** or **motivation** with an AR application • *Usability testing: positive results and issues* Research methods include a set of qualitative, quantitative, and mixed methods. Codes include: • **Quantitative** (survey) • **Qualitative** (interview) • **Mixed methods** (quasi-experiment, case study)	• **Learning effectiveness:** learning gains on word spelling scores (Pu & Zhong, 2018) or self-rated perceived usefulness (Zainuddin & Idrus, 2016) • **Affective outcomes:** motivation, attitude, learning interests (Mahadzir & Phung, 2013) • **Usability testing:** (Papadaki et al., 2013) • **Qualitative:** (Dalim et al., 2016) • **Mixed methods:** (Pu & Zhong, 2018) • **Quantitative:** (Lee et al., 2017)
	Evaluation outcomes	Research results include the main research findings such as affective outcomes, learning outcomes and usability results of the study. Codes includes: • **Learning outcomes:** achieve learning gains and *perceived usefulness* • **Affective outcomes:** motivation and interests • *Usability results: positive results and issues*	• Positive learning gains with AR instructions (A. Ibrahim et al., 2018) • Increased learning motivation with AR instructions (Özcan et al., 2017) • Usability issues (Lee et al., 2017)

learning achievements of learners with an AR instruction and a traditional instruction. Pu and Zhong (2018) reported a study with eight 4- to 8-year-old EFL children who either received an AR instruction or a traditional card-based instruction for 20 minutes in learning English words. The children were tested before and after the instruction on a series of yes/no questions (matching a picture and a given word (e.g., *dog* and *cat*). Juan et al. (2010) presented a study with 32 EFL children (aged 5–6 years) learning Spanish. The children were assigned into one of two groups. Children in Group 1 first received a 15-minute AR instruction and then a 10-minute traditional card-based instruction, while children in Group 2 first received the traditional instruction and then the AR instruction. After each instruction, the children were asked to complete a questionnaire to self report their learning motivation and perceived usefulness of the instruction. R. W. Chen and Chan (2019) presented a study compare AR instruction with a traditional flashcard instruction with 98 EFL children (aged 5–6 years) learning English. The children were divided into two groups and received a 35-minute instruction each week for four weeks. The results of the three studies showed that children had similar learning gains with the two approaches, but with a higher learning motivation with the AR instruction.

Thirty-one studies examined possible effects of AR instruction on learners' affective aspects such as motivation and engagement. All the studies suggested that AR instructions could lead to a higher learning motivation, although one study (Li et al., 2015) reported that learners' motivation level could decrease toward the end of the learning process. Two articles (Safar et al., 2017; Solak & Cakir, 2015) specifically pointed out that learners' greater motivation may lead to better learning outcomes.

Eleven studies reported positive usability testing results with AR applications; however, a few studies also mentioned several issues of using AR applications as instructional tools. The most reported issue was the unstable marker tracking problem with AR applications (Chang et al., 2011; Lee et al., 2017; Luna et al., 2018; Rambli et al., 2013). This instability could be due to inappropriate marker design or inappropriate interaction design (e.g., children's hands blocked markers during interaction). The other problems reported included slow marker detection response time (Lee et al., 2017) and the influence of mobile phones on children's attention and eyesight (He et al., 2014).

Discussion

Effective AR Design and Instructional Strategies

Although almost all the studies showed positive results of learning with AR applications, we noticed three main combined design and instructional strategies that seemed to be effective in supporting early language learning. This became clearer when we only examined the empirical studies that included a matched

control group receiving a traditional instruction and used pre–post tests to evaluate learners' learning gains (i.e., not learners' self-reported perceived learning gains) in their research design (Table 6).

First, the multimedia (and game) design strategies combined with the presentation strategy may support learners' learning gains. For example, in the study presented by Safar et al. (2017), 42 kindergarten children in Kuwait either received an AR instruction or a traditional instruction with teachers for 20 minutes a week for 7 weeks. In the experimental group, an AR application that provides augmented 3D animations, sounds, and words on 2D flashcards was used as a research instrument (*3D multimedia content*). Teachers used this application as an instructional tool to teach the entire classroom of students the English alphabet (*presentation strategy*). The results demonstrated that the children in the AR group had a higher degree of interaction and better learning gains in alphabet tests compared to those in the control group. In another study, Portuguese children used an AR matching game called *MOM* to learn English in class (Barreira et al., 2012). The main activity in MOM was to correctly match a picture card to a word card, in order to view the augmented *multimedia content*, including 3D object, sound, and game feedback (*3D multimedia contents* and *game*). In the instructional process, the children first received an English class given by a teacher, and then started to use the AR application to practice the learnt words (*presentation strategy*). The evaluation results with 26 children (aged 7–9 years) showed that the children who used the MOM game had a superior English learning progress than those who received the traditional instruction. Positive AR learning gains were also found in a study with 60 young adults who learned Ottoman Turkish, the Turkish written in Arabic alphabetic (Özcan et al., 2017) and a study with 20 Chinese kindergarten children learning English words (Y. Chen et al., 2017).

Table 6. Combinations of AR Design and Instructional Strategies.

Types of learning activates	Multimedia contents	Hands-on with physical learning materials	Gamification	Instructional strategies		
				Presentation strategy	Discovery strategy	Collaborative learning
Paper-based word visualization activity	x	X (2D flashcards)	x[a] (matching /puzzle games)	x		x
Location-based word visualization activity and/or spelling games	x	x[a] (3D everyday items)	x (collecting games)		x	x
Augmented spelling games	x	x (2D/3D letters)	x (drills and practice)	xa	x	x

[a]Only a small proportion of the applications in this category contains this particular feature.

Second, the design strategy of location-based learning contents combined with the discovery strategy seem to increase learners' learning gains and learning motivation. For example, in a study that compared the word learning effectiveness between an AR instruction and traditional flashcard instruction, learners in the AR condition wore a *Microsoft Hololens* that enabled a self-exploration of the augmented word spelling on top of the 3D objects (e.g., book, pencil) located in the real-life environment and heard the word sound (N. Ibrahim & Ali, 2018). The results showed that AR was more effective in supporting short-term and long-term learning than the flashcard approach. Learners also felt that AR was more enjoyable than the flashcard approach. In another study with a hand-held AR application in learning German (Santos et al., 2016), the results showed that the location-based AR applications could lead to learners' better retention of words and improve learners' attention and satisfaction compared to the flashcard approach. Similarly, in a study with a hand-held AR *Pokemon Go* game in learning English prefix, root, and suffix knowledge outdoors, the results showed that AR group had significantly better performances than the control classroom group in learning attitudes, satisfactions, and achievements (Wu, 2019).

Third, the game design strategy combined with the discovery strategy may lead to better learning motivation but not better learning gains. In the study presented by (Pu & Zhong, 2018), eight 4- to 8-year-old Chinese children learned English words either by exploring an AR spelling game (*discovery strategy*) or playing with physical cards. The results demonstrated that although the AR group learners had more learning interests and less cognitive load compared to the traditional card group, their learning achievements were similar. In another similar study with 32 children (aged 5–6 years) either playing an AR spelling game or playing physical cards to learn Spanish, the results indicated no significant differences between the two groups about the perceived usefulness, although more children preferred to use the AR game to learn (Juan et al., 2010).

To sum up, in the first strategy, teachers play a dominant role in the instructional process, which may ensure the instruction quality. The AR application is mainly used as an instructional tool to help teachers to better illustrate language knowledge, with the annotations and visualizations of 3D multimedia content (Safar et al., 2017). According to the Meaningful Learning Theory, teacher-centered practice will be more effective in order to construct the link between learners' previous knowledge and new information through deductive instructions, elaborate unclear points, highlight similarities and differences, and motivate learners to pay attention in the learning process (Akdeniz, 2016; Ausubel, 1963). Due to the importance of social interaction in learning, several apps reviewed even attempted to use an animated cartoon tutor to support children's AR learning (Hsieh et al., 2014; Papadaki et al., 2013). This may be explained by Social Agency Theory that posits using verbal and visual cues in computational systems can foster the development of a partnership since learners would

consider human–computer interaction as human–human interaction (Moreno et al., 2001) and trust the computer's instructions.

In the second strategy, location-based applications usually require learners to explore a specific real-world context and construct the relation between the knowledge of words and sentences and the real world. Situated cognition argues that learning cannot be abstracted from the activity, context and culture from which the knowledge was developed (Brown et al., 1989).This is particularly obvious when learning languages that are heavily embedded in contexts and culture. The contextual visualization emphasizes learners' active construction during the learning process, which may benefit children's memorization of words when leveraging the use of the physical environment as the ubiquitous informational cues (Santos et al., 2014).

In the third strategy, learners mainly learn from games with scaffoldings and feedback. The game may promote learners' learning motivation and enable them to learn from trial-and-error exploration. However, previous research has suggested no extra learning gains with such an approach in young children (Juan et al., 2010; Pu & Zhong, 2018). This may be because young children's attention may not be always on the *correct place* without the explicit guidance (R. W. Chen & Chan, 2019). Admittedly, most studies only evaluated short-term learning gains with a small sample of participants with various ages and, with limited samples, we could not make a strong claim about the effectiveness of these strategies. Future research could be done to provide more empirical evidence about their effectiveness.

AR Design Considerations and Opportunities

Leveraging the Unique Affordances of AR in Supporting Early Language Learning. While many articles acknowledged that AR applications enabled multimedia content, allowed for game-based learning, and increased learners' motivation, these benefits also existed in traditional GUIs (Vernadakis et al., 2005; X. Yang et al., 2018) or similarly TUIs (Fan et al., 2016). Therefore, it is necessary to revisit the question: *What unique affordances do AR applications offer for early language learning*?

Previous research has suggested many affordances of AR applications in general educational contexts (Radu, 2014; Santos et al., 2014; Wu et al., 2013). We reflected on their conclusions and related them to review results and identified five main unique AR affordances in supporting early language learning compared to other computational approaches. All five affordances distinguish AR applications from GUIs, but several affordances (A2–A5, Table 7) may also exist in TUIs. Despite this, the specific interactions and implementations in TUIs are different than in AR applications.

1. Transforming abstract language symbols on physical learning materials (e.g., letters, flashcards, objects) to concrete and vivid 2D/3D augmented visual

Table 7. Unique Affordances of AR Interfaces Compared to GUIs and TUIs.

Affordance	AR affordances in general educational contexts	AR affordances and examples in early language learning (AR affordances)	Differences compared to GUIs	Differences compared to TUIs
A1	• Real world annotation (Santos et al., 2014) • Transforming difficult concepts easier to understand (Radu, 2014) • Making the abstract and symbolic concept becoming more concrete ones (Tekedere & Göke, 2016)	Transforming abstract letter/word symbols on physical learning materials to concrete 2D/3D augmented visual representations and auditory sounds	Using concrete 2D/3D visual representations and auditory sounds to demonstrate word spellings, sounds, and meaning.	Transforming abstract letter/word symbols on physical learning materials to concrete 2D/3D visual representations and auditory sounds
A2	• Contextual visualization (Santos et al., 2014) • Presenting relevant educational information at the appropriate time and place (Radu, 2014) • Contextualization (Wu et al., 2013)	Presenting letter/word knowledge relating to everyday objects or real-world locations in an authentic learning environment using off-the-shelf technology	Enabling a digital and simulated learning environments	Presenting letter/word knowledge relating to everyday objects or real-world locations in an authentic learning environment using complex hardware
A3	• Visual-haptic interaction (Santos et al., 2014) • Enabling learners to physically enact the educational concepts (Radu, 2014)	Supporting a variety of hands-on interactions with physical learning materials while viewing the real-time	Usually only enabling touch or mouse/keyboard control	Supporting a variety of hands-on interactions with physical learning materials and while viewing the real-time digital feedback

Table 7. Continued

Affordance	AR affordances in general educational contexts	AR affordances and examples in early language learning		Differences compared to GUIs	Differences compared to TUIs
		AR affordances	augmented digital feedback and actions		
A4	• Directing learners' attention to important educational aspects (Radu, 2014)	Drawing learners' attention to important phonological knowledge using augmented overlays		Drawing learners' attention to important phonological knowledge using digital cues	Drawing learners' attention to important phonological knowledge using digital cues (triggered by actions)
A5	• Permitting learners to interact with spatially challenging phenomena (Radu, 2014)	Enabling word spelling or learning proposition words in a mixed reality space		Only containing 2D spatial interfaces and usually not good for spatial-related tasks	Enabling word spelling or learning proposition words

Note. AR = augmented reality; GUI = graphic user interface; TUI = tangible user interface.

representations and auditory sounds. Compared to multimedia content offered by GUIs (digital–digital), AR applications usually incorporate physical-digital correspondences. The interactions (e.g., grasping, lining up, rotating) on physical learning materials together with the concrete 3D augmented feedback may help learners to concentrate, to understand the abstract language meanings, and to relate the learnt language knowledge to everyday learning tools such as flashcards. The potential learning effects may be explained by the Connectionist Model which suggests using multiple cues (e.g., letters, sounds, pictures) may help with word memorization (Ehri, 2014). Although TUIs also incorporated physical representations, the associated digital contents are usually not directly augmented on top of the physical representations (if there are a few, e.g., *Siftables* (Hunter et al., 2010), they are not in 3D forms augmented in the reality). Learners may need additional cognitive resources to process the transformation between digital information and physical representations based on the Cognitive Load Theory (Price et al., 2008; Sweller, 1988).

2. Presenting language knowledge relating to everyday objects or real-world locations in an authentic learning environment. Situated cognition argues that knowledge cannot be abstracted from the situation from which it was learned (Brown et al., 1989). By teaching words in an authentic environment, learners may have a better chance to use the words in the real-life context. The environments may also serve as a cue to help learners better memorize the words when the connection is created between the two (since learners may unintentionally see certain everyday objects several times a day). TUIs can also associate digital information with physical objects, but they usually require complex hardware embedded with electronics.

3. Enabling a variety of hands-on interactions on virtual objects or physical learning materials with physical affordances (e.g., cubes with physical notches) while viewing real-time augmented feedback on actions through displays (e.g., seeing the word animation when certain cubes are placed in a sequential order). Embodied cognition emphasizes the importance of leveraging the *meaningful* perceptions and actions in learning (Antle, 2013). Learners may use multiple senses such as visual, auditory, tactile senses to construct concrete meanings of abstract concepts; they may also tend to use body actions (e.g., languages, gestures, movements) to interpret concepts based on metaphors. Physical learning materials contain rich multisensory information (e.g., letter shape, color, textures) that can act as explicit or implicit affordances to support learners' language learning. Physical letter shapes with hard edges enable the letter-tracing activity which is suggested as important for at-risk learners (Dehaene, 2009; Kelly & Phillips, 2011). The manipulation of learning materials may also benefit learners in word comprehension by taking advantage of metaphors (Glenberg et al., 2007; Sylla et al., 2016). TUIs also enable hands-on interaction, but there are two

differences compared to AR applications: (a) learners in AR applications can directly interact with either virtual objects or physical learning materials while interactions in most TUIs are directly on physical learning materials; and (b) learners in AR applications can view real-time feedback and hands-on actions simultaneously through the display while TUIs cannot.

4. Drawing learners' attention to important phonological knowledge using augmented overlay. Mayer's Cognitive Theory of Multimedia Learning suggests three main characteristics in the cognitive process of human beings: *visual-auditory dual channels*, *limited capacity*, and *active processing* (Mayer, 2014). The limited capacity mainly posits that learners are limited in the amount of information that can be processed; therefore, they need to select relevant information to ensure meaningful learning occurs. In addition to offering educational information, the augmented overlay can also serve as an attentional cue to better draw learners' attention to language learning (e.g., by using colors or animations). Although GUIs and TUIs also enable this feature, GUIs do not associate the cues with physical actions (Kast et al., 2007) and TUIs may require complex hardware to achieve this (Antle et al., 2015).

5. Supporting word spelling or proposition word learning in both physical and digital space.[6] By leveraging the use of congruent spatial mappings between physical and digital space in AR applications, learners can perform spelling tasks in either physical space (e.g., letters on the table) or digital space (e.g., digital letters on the screen) by fully utilizing their embodied knowledge (e.g., left or right). This feature is particularly beneficial for learning proposition words when using head-mounted displays or handheld devices (not desktop computers) so that learners can directly see an augmented virtual object *on the left or on the right* of a certain real-world physical object and hear the associated sound or receive other digital feedback simultaneously. Despite the congruent spatial mappings, the associated digital virtual objects in TUIs are usually in a separated display (see representative examples of commercial tangible language learning applications with an iPad, *Osmo*[7] or *Marbotics*[8]) rather than in a 3D mixed reality space.

The affordances first can be used to identify missing opportunities in the current design space. For example, our review showed a significant proportion of the AR applications utilized augmented content to visualize word information (A1) rather than to direct learners' attention (A4); and fewer applications fully leveraged the contextual information (e.g., locations) to support language learning (A2). The affordances also can be used to provide justifications for future designs (A3). For example, the A1 to A2 may be particularly beneficial for EFL learners who do not have proficient oral skills and lack contextual knowledge of the language. The A3 to A4 could be used to help learners with attentional problems in language learning. The A5 may help learners to better leverage the use of embodied knowledge to learn syllable/word construction

(e.g., physically placing letters in a sequential order, grouping letter pairs together or dividing syllables separately).

Emphasizing on Phonological Knowledge Rather Than Whole Word Reading. Our review also showed most AR applications focused on visualizations of whole words rather than phonological knowledge. Phonological knowledge is extremely important in early literacy instruction, particularly for children at-risk for reading difficulties (Vellutino et al., 2004). According to Ehri (2014), word spelling can be further divided into four stages, including *prealphabetic*, *partial alphabetic*, *full alphabetic*, and *consolidated alphabetic* stages. There are a number of activities that are commonly used in school education to teach learners phonological knowledge in each stage. For example, *letter names and letter sounds* that focus on single alphabet phonetic training is used in *prealphabetic stage* while activities on *onset and rime*, *letter insertion, omission*, or *substitution* can be used in *partial alphabetic stage*, and *word decoding* and *the alphabetic principle* can be taught to learners in *full and consolidated alphabetic* stages to help them gradually develop solid word spelling ability (Kelly & Phillips, 2011; Rello et al., 2012). The variety of phonological learning activities could be rich resources for future AR applications.

In addition, designers can use the five affordances as guidance to consider how to better support the learning of phonological knowledge. For example, A4 suggests the potential of using augmented overlay as an attentional cue (rather than purely displaying knowledge such as letters, sounds or pictures). To draw learners' attention to the decoding process, augmented letter animations could be used to visualize the letter *combining* or *separating* actions during word decoding processes; augmented color flash could be used to illustrate grouped letter patterns; single phonological sounds (*cat*-/kat/, *flag*-/fl-a-ɡ/) could also be designed to promote learners' phonological awareness. A2 suggests the promise of leveraging location-relevant information to teach language knowledge. In this case, designers could design location-based games wherein learners are required to collect all the words starting with a certain letter sound (e.g., *book*, *bag*, *ball* all containing/b/sound) or fit a certain alphabetic rule (e.g., *book*, *hook*) in the environment.

Considering the Use of AR Applications With Teachers and Learners in Context. Almost two thirds of the studies included teachers during the AR instructional process. Therefore, we advocate that designers consider the practical factors of teachers and learners using AR applications together from their initial design stage. Specifically, our review suggests the following six aspects that should be given attention.

1. *Providing effective blended instructions with AR applications*: Traditional instructional methods and resources that have been tested for decades have

shown their effectiveness. The aim of AR instruction is to *enhance* NOT *replace* traditional instruction. Therefore, designers should consider how to integrate AR instruction into traditional instruction practices and contexts to provide effective blended learning opportunities. Our review suggests that AR instruction can (a) start from the instruction phase where the AR applications are commonly used as *instructional tools* for knowledge demonstration (Solak & Cakir, 2015), (b) start at the practice phase where the AR applications are usually game-based activities (Barreira et al., 2012), or (c) cover both instruction and practice phases by containing multiple modes to support teachers/learners switching between phases (Fan et al., 2018; Papadaki et al., 2013; Rambli et al., 2013). We suggest that designers consider these issues by taking account of the specific learning goal (e.g., phonological awareness, the alphabetic principle, vocabularies), the AR learning activity, and learning contexts.

2. *Considering teachers' role and level of participation in an AR instruction*: Our review has revealed four main types of teachers' roles, including *a dominate instructor* (Solak & Cakir, 2015), *an activity facilitator* (Sytwu & Wang, 2016; Vate-U-Lan, 2012), *an observer* (Papadaki et al., 2013; Rambli et al., 2013), and *a mixed role* (Barreira et al., 2012; Fan et al., 2018). The level of participation is decreased from the *instructor* to *observer*, according to the changes of the activities, for example, from instruction to practice. Therefore, on one hand, we suggest designers enabling teachers to switch roles during various AR instruction phases; and, on the other hand, we advocate to enhance teachers' level of participation while acting as a *dominant instructor*, by providing more opportunities for teachers to control and adjust the use of AR applications (e.g., adding words, sounds, or pictures).

3. *Splitting the scaffoldings between an AR application and a teacher*: We found most AR applications reviewed contained various levels of scaffolding to support learners, including (a) scaffolding that *enforced* corrective actions, for example, in the spelling game in AR Magic English, the digital letter would be restored to its original location if it is wrong (Pu & Zhong, 2018); and (b) scaffolding that provided partial/full answers to *encourage* corrective actions, for example, providing three levels of hints which repeat the task question, provide partial answer and full answer (Fan et al., 2018). However, what seems unclear here is whether the help provided by AR applications themselves is sufficient, and if not, *when* and *what* kinds of scaffolding should be provided by teachers and supported by application design. We suggest designers consider these aspects in their design.

4. *Designing strategies to support positive collaborative learning*: Although several articles contained collaborative learning activities, most of the collaborative learning activities were limited to sitting and using the AR application together (Barreira et al., 2012; Vate-U-Lan, 2012) or holding a shared tablet to find certain real-world objects and words (Sytwu & Wang, 2016), from

which we still could not know how much collaborative learning actually occurred and how equally each learner participated in the learning activity. Previous research has suggested the importance of equal verbal and non-verbal participation in a collaborative learning activity (Antle et al., 2014). We encourage designers leveraging various design strategies (e.g., jigsaw methods) and investigating whether and how these design strategies could support equal and in-depth discussion and effective physical interaction among students as well as the potential influence on students' learning outcomes.

5. *Exploring AR design in various learning contexts*: We also noticed a few articles attempted to cover the personalized use of AR applications, for example, learning styles (Hsu, 2017), and the uses of AR applications in various learning contexts (e.g., individual, small groups, classroom; Vate-U-Lan, 2012) and learning groups (e.g., readers with attention deficit hyperactivity disorder and Autism; Lin et al., 2016; Tang et al., 2019). Therefore, one potential design direction is to explore the specific learning activities and design functions an AR application should have to best facilitate the effective learning for specific learning contexts. And one particular issue related to location-based AR applications is to consider what affordances and constraints of the learning environment may impact learning, for example, location-relevant learning objects and safety issues when walking around in a class.

6. *Offering an easy access for teachers to update AR contents:* One major issue revealed in previous reviews was the inflexibility of updating AR content for teachers (Bacca et al., 2014; Santos et al., 2014; Wu et al., 2013). This problem is particularly obvious in early language learning contexts since there could be numerous words to learn in the database. We noticed researchers in a few articles attempted to utilize easy-to-learn commercial AR tools, such as *Aurasm* and *Zoobust*, to produce easy-to-update AR content. However, because of the limited functions of the tools, these AR applications only allow users to add pictures/texts to physical cards. The best applications, produced by professional software (e.g., *Unity* and *Vuforia*), however, did not describe the specific plans for extending learning content or making it changeable by teachers. We suggest designers create lessons according to teachers' specific requirements, as well as providing a user-friendly authoring tool (e.g., a database sheet for adding words/pictures) so teachers can update learning materials as their needs change.

Future Evaluation Directions

Although the majority of the articles showed positive learning gains, most of studies only evaluated short-term learning gains on word memorizations with a particular AR application without any compared counterpart. Therefore, we

suggest the potential research directions on the evaluation of associated learning effects with a specific design strategy and the long-term learning outcomes on both the memorization and the application of the language knowledge in applied contexts.

Understanding the Learning Effect of Various Design Strategies. Using multiple verbal and non-verbal digital representations may help learners with memorization based on Connectionist Model (Ehri, 2014) and Dual Coding Theory (Clark & Paivio, 1991), but they may also require more cognitive load from learners (Dunleavy et al., 2009). It is necessary to investigate how various multimedia design strategies (e.g., using 3D models or animation, color cues) influence learners' learning gains specifically. We also noticed that 3D models were often used to visualize concrete words rather than abstract words. It would be interesting to explore the effective design strategies of multimedia in AR applications to support the instruction of abstract words (e.g., *bad*, *thin*).

Moreover, previous research suggested that it is more appropriate to use lowercase letters in early language learning, especially for at-risk learners (Dehaene, 2009; Fan et al., 2016). However, we noticed that the use of uppercase letters and lowercase letters varied among the reviewed AR applications. Therefore, it is important to understand the potential impacts of letter choices such as letter type, font, size, and augmented perspective (e.g., 90° or 45° based on the ground) on learners' language learning outcomes in future research.

Furthermore, our review showed most AR applications used gamification, but it is still unclear how game strategies impact children's learning gains. In addition, the current game genres in AR applications are still very limited, with most of them coming in the form of *quizzes*. Hung et al. (2018) identified eight game genres and strategies of digital games for language learning and several strategies (e.g., *narrative* and *role play*) that could be integrated with AR applications. We believe that it is crucial to understand how various game strategies such as virtual tutors, role play, or narrative influence learners' learning experience, motivation, and outcomes.

Investigating the Generalization and Maintenance of AR Learning Gains As Well As the Effectiveness of Instructions With AR and Similar Computational Systems. Similar to the results found in previous reviews on AR learning games (Li et al., 2017), the most frequently used measurement of learning achievement was a pretest and posttest on vocabulary knowledge. However, the pretest and posttest measurement method often only tested whether learners could *remember* learnt literacy knowledge (e.g., remembering the trained vocabulary) rather than whether learners could *apply* the language knowledge (e.g., using the learnt rules/letters to spell out similar new words or use the learnt words in contexts). Future studies could focus on more complex cognitive processes to investigate whether learners could apply the language knowledge.

The majority of the studies used quantitative methods to examine learners' learning gains. In these studies, learning scores were collected and compared to investigate learners' achievements. However, almost all the studies measured the learning outcomes immediately after the use of the AR applications. Only one study measured the follow-up learning gains. And most studies only included a small sample size. Because most of the learners never used AR applications before there is a potential novelty effect impacting the research findings. Therefore, future research could focus on both short-term and long-term learning gains and incorporate larger sample sizes. We encourage investigations into whether children could maintain their learning achievements over time. In addition, longitudinal case studies could be conducted to better understand the contextual use of AR applications with teachers/parents and learners in applied settings like the classroom or at home.

We found that more than half of the studies did not contain any control groups when evaluating AR applications. Therefore, it is difficult to know the effectiveness of AR instruction in supporting learners' learning gains compared to traditional approaches or any other computational approaches such as instructions with desktop applications or tangible systems. Future research could investigate how different computational instruction applications impact learners' motivation and learning gains. This may also help to identify what specific design features of AR applications combined with specific instructional strategies may benefit language learning.

Conclusion

This review presents AR as a promising technology for supporting early language learning. The major results can be summarized as follows: (a) we identified three types of AR learning activities: augmented word spelling games, card-based word visualization activity, and location-based word annotations activity and/or spelling games; (b) the preferred design strategies included 3D multimedia content with sounds, interaction with 2D physical flashcards, gamification, and location-based vocabulary learning; (c) most AR applications used discovery instructional strategy, followed by collaborative and presentation strategies; and (d) learning effectiveness, motivation, and system usability were often evaluated in the studies with mixed methods or quantitative methods.

We suggest a set of design considerations and design opportunities for future design of AR applications for early language learning: (a) 3D multimedia content with presentation strategy and location-based design with the discovery strategy seem to be effective in supporting learning, while game design with the discovery strategy seem to benefit learning motivation. (b) Designers should leverage the use of unique affordances of AR to support language learning, including (i) transforming abstract language symbols on physical learning materials to concrete and vivid 2D/3D augmented visual representations and

auditory sounds; (ii) presenting language knowledge relating to everyday objects or real-world locations in an authentic learning environment; (iii) enabling a variety of hands-on interactions on virtual objects or physical learning materials with physical affordances; (iv) drawing learners' attention to important phonological knowledge using augmented overlay; and (v) supporting word spelling or preposition word learning in physical and digital space. (c) Learning goals should pay attention to phonological knowledge which is crucial to early language learning; (d) Designers should consider the use context of AR in the design process, including (i) providing a smooth transition between traditional instruction and AR instruction, (ii) considering teachers' role and level of participation in AR instruction, (iii) splitting roles of scaffolding between an AR application and a teacher, (iv) designing strategies to support positive collaborative learning, (v) exploring AR design in specific learning contexts, and (vi) offering easy access for teachers to update AR content. We also suggest two research directions: (a) understanding the learning effect of various design strategies and (b) investigating the generalization and maintenance of AR learning gains as well as the effectiveness of instruction with AR applications and similar computational systems.

Declaration of Conflicting Interests

Funding

ORCID iD

Min Fan https://orcid.org/0000-0001-5323-8168

Supplemental material

Supplemental material for this article is available online.

Notes

1. GUIs are a standard computing paradigm which graphical representations displayed on the screen are manipulated with generic remote controllers such as mice and keyboard (Ishii, 2008).
2. TUIs are a computing paradigm where the real world is augmented by embedded physical objects with digital information (Ishii & Ullmer, 1997).

3. A *narrative review* uses a qualitative interpretation of the literature (H. Yang & Tate, 2009).
4. A *systematic review* systematically searches, filters, and classifies articles in order to produce a qualitative and quantitative summary (H. Yang & Tate, 2009).
5. A *meta-review* aims at statistically providing support by analysing the quantitative results of many empirical studies (H. Yang & Tate, 2009).
6. This affordance is kindly related to Affordances (1) and (3), but focuses more on the spatial benefit offered by AR applications.
7. See http://playosmo.com
8. See http://www.marbotic.fr

References

Akçayır, M., & Akçayır, G. (2017). Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educational Research Review*, *20*, 1–11. https://doi.org/10.1016/j.edurev.2016.11.002

Akdeniz, C. (2016). *Instructional process and concepts. In theory and practice: Improving the teaching process* (1st ed.). Springer Nature. https://link.springer.com/book/10.1007%2F978-981-10-2519-8

Antle, A. N. (2013). Research opportunities: Embodied child–computer interaction. *International Journal of Child-Computer Interaction*, *1*(1), 30–36. https://doi.org/10.1016/j.ijcci.2012.08.001

Antle, A. N., Fan, M., & Cramer, E. S. (2015). *PhonoBlocks: A tangible system for supporting dyslexic children learning to read* [Paper presentation]. 9th International Conference on Tangible, Embedded, and Embodied Interaction (pp. 533–538). ACM. https://doi.org/10.1145/2677199.2687897

Antle, A. N., Warren, J. L., May, A., Fan, M., & Wise, A. F. (2014). *Emergent dialogue: Eliciting values during children's collaboration with a tabletop game for change* [Paper presentation]. Proceedings of the 13th Conference on Interaction Design and Children (pp. 37–46). Association for Computing Machinery. https://doi.org/10.1145/2593968.2593971

Apel, K. (2011). What is orthographic knowledge?. *Language, Speech, and Hearing Services in Schools*, *42*(4), 592–603. https://doi.org/10.1044/0161-1461(2011/10-0085)

Ati, M., Kabir, K., Abdullahi, H., & Ahmed, M. (2018). *Augmented reality enhanced computer aided learning for young children* [Paper presentation]. 2018 IEEE Symposium on Computer Applications Industrial Electronics (pp. 129–133). IEEE. https://doi.org/10.1109/ISCAIE.2018.8405457

Ausubel, D. P. (1963). *The psychology of meaningful verbal learning*. Grune & Stratton. https://psycnet.apa.org/record/1964-10399-000

Azuma, R. T. (1997). A survey of augmented reality. *Presence: Teleoperators & Virtual Environments*, *6*(4), 355–385. https://doi.org/10.1162/pres.1997.6.4.355

Bacca, J., Baldiris, S., Fabregat, R., & Graf, S. (2014). Augmented reality trends in education: A systematic review of research and applications. *Educational Technology & Society*, *17*(4), 133–149. https://www.jstor.org/stable/jeductechsoci.17.4.133

Barreira, J., Bessa, M., Pereira, L. C., Adão, T., Peres, E., & Magalhães, L. (2012). *MOW: Augmented Reality game to learn words in different languages: Case study: Learning English names of animals in elementary school* [Paper presentation]. 7th

Iberian Conference on Information Systems and Technologies (pp. 1–6). IEEE. https://ieeexplore.ieee.org/abstract/document/6263236

Bhadra, A., Brown, J., Ke, H., Liu, C., Shin, E., Wang, X., & Kobsa, A. (2016). *ABC3D—Using an augmented reality mobile game to enhance literacy in early childhood* [Paper presentation]. 2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops) (pp. 1–4). IEEE. https://doi.org/10.1109/PERCOMW.2016.7457067

Boonbrahm, S., Kaewrat, C., & Boonbrahm, P. (2015). *Using augmented reality technology in assisting English learning for primary school students* [Paper presentation]. Learning and Collaboration Technologies (pp. 24–32). Springer. https://doi.org/10.1007/978-3-319-20609-7_3

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, *18*(1), 32–42. https://doi.org/10.3102/0013189X018001032

Chang, Y. J., Chen, C. H., Huang, W. T., & Huang, W. S. (2011). *Investigating students' perceived satisfaction, behavioral intention, and effectiveness of English learning using augmented* reality [Paper presentation]. 2011 IEEE International Conference on Multimedia and Expo (pp. 1–6). IEEE. 10.1109/ICME.2011.6012177

Chen, C., & Wang, C. (2015). *The effects of learning style on mobile augmented-reality-facilitated English vocabulary learning* [Paper presentation]. 2015 2nd International Conference on Information Science and Security (pp. 1–4). IEEE. https://doi.org/10.1109/ICISSEC.2015.7371036

Chen, R. W., & Chan, K. K. (2019). Using augmented reality flashcards to learn vocabulary in early childhood education. *Journal of Educational Computing Research*, *57*(7), 1812–1831. https://doi.org/10.1177/0735633119854028

Chen, Y., Zhou, D., Wang, Y., & Yu, J. (2017). Application of augmented reality for early childhood English teaching [Paper presentation]. 2017 International Symposium on Educational Technology (pp. 111–115). IEEE. https://doi.org/10.1109/ISET.2017.34

Cheng, K. H., & Tsai, C. C. (2013). Affordances of augmented reality in science learning: Suggestions for future research. *Journal of Science Education and Technology*, *22*(4), 449–462. https://doi.org/10.1007/s10956-012-9405-9

Cieza, E., & Lujan, D. (2018). Educational mobile application of augmented reality based on markers to improve the learning of vowel usage and numbers for children of a kindergarten in Trujillo. *Procedia Computer Science*, *130*, 352–358. https://doi.org/10.1016/j.procs.2018.04.051

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, *3*(3), 149–210. https://doi.org/10.1007/BF01320076

Colwell, J., & Hutchison, A. C. (2015). Supporting teachers in integrating digital technology into language arts instruction to promote literacy. *Journal of Digital Learning in Teacher Education*, *31*(2), 56–63. https://doi.org/10.1080/21532974.2014.991813

Dalim, C. S. C., Piumsomboon, T., Dey, A., Billinghurst, M., & Sunar, S. (2016). TeachAR: An interactive augmented reality tool for teaching basic English to non-native children [Paper presentation]. 2016 IEEE International Symposium on Mixed and Augmented Reality (pp. 344–345). IEEE. https://doi.org/10.1109/ISMAR-Adjunct.2016.0113

Dehaene, S. (2009). *Reading in the brain: The new science of how we read* (1st ed.). Penguin Group.

Dunleavy, M., Dede, C., & Mitchell, R. (2009). Affordances and limitations of immersive participatory augmented reality simulations for teaching and learning. *Journal of Science Education and Technology*, *18*(1), 7–22. https://doi.org/10.1007/s10956-008-9119-1

Earnshaw, A., & Seargeant, A. (2005). *Dealing with dyslexia and other reading difficulties*. Pearson/Prentice Hall.

Ehri, L. C. (2014). Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Scientific Studies of Reading*, *18*(1), 5–21. https://doi.org/10.1080/10888438.2013.819356

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing*, *62*(1), 107–115. https://doi.org/10.1111/j.1365-2648.2007.04569.x

Fan, M., Antle, A. N., & Cramer, E. S. (2016). Design rationale: Opportunities and recommendations for tangible reading systems for children [Paper presentation]. 15th International Conference on Interaction Design and Children (pp. 101–112). ACM. https://doi.org/10.1145/2930674.2930690

Fan, M., Antle, A. N., Hoskyn, M., Neustaedter, C., & Cramer, E. S. (2017). Why tangibility matters: A design case study of at-risk children learning to read and spell [Paper presentation]. 2017 Conference on Human Factors in Computing Systems (pp. 1805–1816). ACM Press. https://doi.org/10.1145/3025453.3026048

Fan, M., Antle, A. N., & Sarker, S. (2018). From tangible to augmented: Designing a PhonoBlocks reading system using everyday technologies [Paper presentation]. Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems (pp. LBW555:1–LBW555:6). ACM Press. https://doi.org/10.1145/3170427.3188459

Fan, M., Baishya, U., Mclaren, E.S., Antle, A. N., Sarker, S., & Vincent, A. (2018a). Block talks: A tangible and augmented reality toolkit for children to learn sentence construction [Paper presentation]. Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems (pp. LBW056:1–LBW056:6). ACM Press. https://doi.org/10.1145/3170427.3188576

Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. Patterson, J. C. Marshall, & M. Coltheart (Eds.), Surface dyslexia: Neuropsychological and cognitive studies of phonological reading (pp. 301–330). Erlbaum.

Glenberg, A. M., Brown, M., & Levin, J. R. (2007). Enhancing comprehension in small reading groups using a manipulation strategy. *Contemporary Educational Psychology*, *32*(3), 389–399. https://doi.org/10.1016/j.cedpsych.2006.03.001

Goh, W. B., Kasun, L. L. C., Fitriani, Tan, J., & Shou, W. (2012). The i-Cube: Design considerations for block-based digital manipulatives and their applications [Paper presentation]. Designing Interactive Systems Conference (pp. 398–407). ACM. https://doi.org/10.1145/2317956.2318016

He, J., Ren, J., Zhu, G., Cai, S., & Chen, G. (2014). Mobile-based AR application helps to promote EFL children's vocabulary study [Paper presentation]. The 14th International Conference on Advanced Learning Technologies (pp. 431–433). IEEE. https://doi.org/10.1109/ICALT.2014.129

Hengeveld, B., Hummels, C., van Balkom, H., Voort, R., & de Moor, J. (2013). Wrapping up LinguaBytes, for now [Paper presentation]. 7th International Conference on Tangible, Embedded and Embodied Interaction (pp. 237–244). ACM. https://doi.org/10.1145/2460625.2460664

Ho, S. C., Hsieh, S. W., Sun, P. C., & Chen, C. M. (2017). To activate English learning: Listen and speak in real life context with an AR featured u-learning system. *Journal of Educational Technology & Society*, *20*(2), 176–187.

Hornecker, E., & Dünser, A. (2007). Supporting early literacy with augmented books-experiences with an exploratory study [Paper presentation]. Proceedings of the German Society of Informatics Annual Conference (pp. 555–559). https://dl.gi.de/handle/20.500.12116/22549

Hsieh, M. C., Kuo, F. R., & Lin, H. C. K. (2014). The effect of employing AR interactive approach on students' English preposition learning performance. *Journal of Computers and Applied Science Education*, *1*, 45–60. http://bit.kuas.edu.tw/~jcase/2014/vol1/JCASE-2014-01-003.pdf

Hsu, T. C. (2017). Learning English with augmented reality: Do learning styles matter? *Computers & Education*, *106*, 137–149. https://doi.org/10.1016/j.compedu.2016.12.007

Hung, H. T., Yang, J. C., Hwang, G. J., Chu, H. C., & Wang, C. C. (2018). A scoping review of research on digital game-based language learning. *Computers & Education*, *126*, 89–104. https://doi.org/10.1016/j.compedu.2018.07.001

Hunter, S., Kalanithi, J., & Merrill, D. (2010). Make a Riddle and TeleStory: Designing children's applications for the Siftables Platform [Paper presentation]. Proceedings of the 9th International Conference on Interaction Design and Children (pp. 206–209). ACM. https://doi.org/10.1145/1810543.1810572

Ibáñez, M.-B., & Delgado-Kloos, C. (2018). Augmented reality for STEM learning: A systematic review. *Computers & Education*, *123*, 109–123. https://doi.org/10.1016/j.compedu.2018.05.002

Ibrahim, A., Huynh, B., Downey, J., Höllerer, T., Chun, D., & O'donovan, J. (2018). Arbis pictus: A study of vocabulary learning with augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, *24*(11), 2867–2874. https://doi.org/10.1109/TVCG.2018.2868568

Ibrahim, N., & Ali, N. M. (2018). A conceptual framework for designing virtual heritage environment for cultural learning. *Journal on Computing and Cultural Heritage*, *11*(2), 11. https://doi.org/10.1145/3117801

Ishii, H. (2008). The tangible user interface and its evolution. *Communications of the ACM*, *51*(6), 32–36. https://dl.acm.org/doi/fullHtml/10.1145/1349026.1349034

Ishii, H. and Ullmer, B. (1997). Tangible bits: towards seamless interfaces between people, bits and atoms. 1997 Conference on Human Factors in Computing Systems (pp. 234–241). ACM Press. https://dl.acm.org/doi/pdf/10.1145/258549.258715

Juan, C. M., Llop, E., Abad, F., & Lluch, J. (2010). Learning words using augmented reality [Paper presentation]. Advanced Learning Technologies, 2010 IEEE (pp. 422–426). IEEE. https://doi.org/10.1109/ICALT.2010.123

Kast, M., Meyer, M., Vögeli, C., Gross, M., & Jäncke, L. (2007). Computer-based multisensory learning in children with developmental dyslexia. *Restorative Neurology and Neuroscience*, *25*(3–4), 355–369. https://content.iospress.com/articles/restorative-neurology-and-neuroscience/rnn253417

Kelly, K., & Phillips, S. (2011). *Teaching literacy to learners with dyslexia: A multisensory approach*. SAGE.

Koutromanos, G., Sofos, A., & Avraamidou, L. (2015). The use of augmented reality games in education: A review of the literature. *Educational Media International*, *52*(4), 253–271. https://doi.org/10.1080/09523987.2015.1125988

Kucuk, S., Yilmaz, R. M., & Goktas, Y. (2014). Augmented reality for learning English: Achievement, attitude and cognitive load levels of students. *Egitim Ve Bilim-Education and Science*, *39*(176), 393–404. https://doi.org/10.15390/EB.2014.3595

Lee, L., Chau, C., Chau, C., & Ng, C. (2017). Using augmented reality to teach kindergarten students English vocabulary [Paper presentation]. 2017 International Symposium on Educational Technology (ISET) (pp. 53–57). IEEE. https://doi.org/10.1109/ISET.2017.20

Li, J., van der Spek, E. D., Feijs, L., Wang, F., & Hu, J. (2017) Augmented reality games for learning: A literature review. In N. Streitz & P. Markopoulos (Eds.), *Distributed, ambient and pervasive interactions* (1st ed., pp. 612–626). Springer. https://doi.org/10.1007/978-3-319-58697-7_46

Li, M. S., Chen, M. Y., Whittinghill, D. M., & Vorvoreanu, M. (2015). A pilot study exploring augmented reality to increase motivation of Chinese college students learning English. *Computers in Education Journal*, *25*(1), 23–33. https://doi.org/10.1007/978-3-319-58697-7_46

Limsukhawat, S., Kaewyoun, S., Wongwatkit, C., & Wongta, J. (2016). A development of augmented reality-supported mobile game application based on jolly phonics approach to enhancing English phonics learning performance of ESL learners. In W. Chen, J. Yang, S. Murthy, S. Wong and S. Iyer (Eds.), Proceedings of the 24th International Conference on Computers in Education (1st ed., pp. 483–488). Asia Pacific Society for Computers in Education.

Lin, C. Y., Yu, W. J., Chen, W. J., Huang, C. W., & Lin, C. C. (2016). The effect of literacy learning via mobile augmented reality for the students with ADHD and reading disabilities [Paper presentation]. International Conference on Universal Access in Human-Computer Interaction (pp. 103–111). Springer. https://doi.org/10.1007/978-3-319-40238-3_11

Luna, J., Treacy, R., Hasegawa, T., Campbell, A., & Mangina, E. (2018). Words worth learning-Augmented literacy content for ADHD students [Paper presentation]. 2018 IEEE Games, Entertainment, Media Conference (pp. 1–9). IEEE. https://doi.org/10.1109/GEM.2018.8516483

Mahadzir, N. N., & Phung, L. F. (2013). The use of augmented reality pop-up book to increase motivation in English language learning for national primary school. *Journal of Research & Method in Education*, *1*(1), 26–38. https://doi.org/10.9790/7388-0112638

Mayer, R. E. (2014). Cognitive theory of multimedia learning. In The Cambridge handbook of multimedia learning (2nd ed., pp. 43–71). Cambridge University Press.

Milgram, P., Takemura, H., Utsumi, A., & Kishino, F. (1995). Augmented reality: A class of displays on the reality-virtuality continuum. *Telemanipulator and Telepresence Technologies*, *2351*, 282–293. https://doi.org/10.1117/12.197321

Mioduser, D., Tur-Kaspa, H., & Leitner, I. (2000). The learning value of computer-based instruction of early reading skills. *Journal of Computer Assisted Learning*, *16*(1), 54–63. https://doi.org/10.1046/j.1365-2729.2000.00115.x

Moreno, R., Mayer, R. E., Spires, H. A., & Lester, J. C. (2001). The case for social agency in computer-based teaching: Do students learn more deeply when they interact

with animated pedagogical agents? *Cognition and Instruction*, *19*(2), 177–213. https://doi.org/10.1207/S1532690XCI1902_02

Nicolson, R., Fawcett, A., & Nicolson, M. (2000). Evaluation of a computer-based reading intervention in infant and junior schools. *Journal of Research in Reading*, *23*(2), 194–209. https://doi.org/10.1111/1467-9817.00114

Özcan, M. F., Özkan, Â., & Sahin, N. (2017). The influence of the augmented reality application on students' performances in Ottoman Turkish readings. *Universal Journal of Educational Research*, *5*, 27–33.

Papadaki, E., Zabulis, X., Ntoa, S., Margetis, G., Koutlemanis, P., Karamaounas, P., & Stephanidis, C. (2013). The book of Ellie: An interactive book for teaching the alphabet to children [Paper presentation]. 2013 IEEE International Conference on Multimedia and Expo Workshops (pp. 1–6). IEEE. https://doi.org/10.1109/ICMEW.2013.6618341

Price, S., Sheridan, J. G., Falcao, T. P., & Roussos, G. (2008). Towards a framework for investigating tangible environments for learning. *International Journal of Arts and Technology*, *1*(3–4), 351–368. https://doi.org/10.1504/IJART.2008.022367

Pu, M., & Zhong, Z. (2018). Development of a situational interaction game for improving preschool children' performance in English-vocabulary learning [Paper presentation]. Proceedings of the 2018 International Conference on Distance Education and Learning (pp. 88–92). ACM. https://doi.org/10.1145/3231848.3231851

Radu, I. (2014). Augmented reality in education: A meta-review and cross-media analysis. *Personal and Ubiquitous Computing*, *18*(6), 1533–1543. https://doi.org/10.1007/s00779-013-0747-y

Rambli, D. R. A., Matcha, W., & Sulaiman, S. (2013). Fun learning with AR alphabet book for preschool children. *Procedia Computer Science*, *25*, 211–219. https://doi.org/10.1016/j.procs.2013.11.026

Rello, L., Bayarri, C., & Gorriz, A. (2012). What is wrong with this word? Dyseggxia: A game for children with dyslexia [Paper presentation]. Proceedings of the 14th International ACM SIGACCESS Conference on Computers and Accessibility (pp. 219–220). ACM. https://doi.org/10.1145/2384916.2384962

Sadoski, M., & Paivio, A. (2013). *Imagery and text: A dual coding theory of reading and writing* (2nd ed.). Routledge.

Safar, A. H., Al-Jafar, A. A., & Al-Yousefi, Z. H. (2017). The effectiveness of using augmented reality apps in teaching the English alphabet to kindergarten children: A case study in the state of Kuwait. *EURASIA Journal of Mathematics*, *Science & Technology Education*, *13*(2), 417–440. https://doi.org/10.12973/eurasia.2017.00624a

Sansosti, F. J., Powell-Smith, K. A., & Kincaid, D. (2004). A research synthesis of social story interventions for children with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities*, *19*(4), 194–204. https://doi.org/10.1177/10883576040190040101

Santos, M. E. C., Chen, A., Taketomi, T., Yamamoto, G., Miyazaki, J., & Kato, H. (2014). Augmented reality learning experiences: Survey of prototype design and evaluation. *IEEE Transactions on Learning Technologies*, *7*(1), 38–56. https://doi.org/10.1109/TLT.2013.37

Santos, M. E. C., Taketomi, T., Yamamoto, G., Rodrigo, M. M. T., Sandor, C., & Kato, H. (2016). Augmented reality as multimedia: The case for situated vocabulary

learning. *Research and Practice in Technology Enhanced Learning*, *11*(1), 4. https://doi.org/10.1186/s41039-016-0028-2

Silva, M., Roberto, R., & Teichrieb, V. (2013). Evaluating an educational system based on projective augmented reality. *Brazilian Symposium on Computers in Education*, *24*(1), 214–223. http://dx.doi.org/10.5753/cbie.sbie.2013.214

Slavin, R. E. (1991). *Student team learning: A practical guide to cooperative learning* (3rd ed.). National Education Association. https://eric.ed.gov/?id = ED339518

Solak, E., & Cakir, R. (2015). Exploring the effect of materials designed with augmented reality on language learners' vocabulary learning. *Journal of Educators Online*, *12*(2), 50–72. https://eric.ed.gov/?id = EJ1068381

Song, S. H., & Keller, J. M. (2001). Effectiveness of motivationally adaptive computer-assisted instruction on the dynamic aspects of motivation. *Educational Technology Research and Development*, *49*(2), 5. https://doi.org/10.1007/BF02504925

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257–285. https://doi.org/10.1016/0364-0213(88)90023-7

Sylla, C., Pereira, I. S. P., Coutinho, C. P., & Branco, P. (2016). Digital manipulatives as scaffolds for preschoolers' language development. *IEEE Transactions on Emerging Topics in Computing*, *4*(3), 439–449. https://doi.org/10.1109/TETC.2015.2500099

Sysoev, I., Hershman, A., Fine, S., Traweek, C., & Roy, D. (2017). SpeechBlocks: A constructionist early literacy app [Paper presentation]. 2017 Conference on Interaction Design and Children (pp. 248–257). ACM. https://doi.org/10.1145/3078072.3079720

Sytwu, T. A., & Wang, C. H. (2016). An investigation of the effects of individual differences on mobile-based augmented reality English vocabulary learning. In D. Churchill, J. Lu, T. Chiu, & B. Fox (Eds.), Mobile learning design (1st ed., pp. 387–410). Springer. https://doi.org/10.1007/978-981-10-0027-0_23

Tang, T. Y., Xu, J., & Winoto, P. (2019). Automatic object recognition in a light-weight augmented reality-based vocabulary learning application for children with autism [Paper presentation]. 3rd International Conference on Innovation in Artificial Intelligence (pp. 65–68). ACM. https://doi.org/10.1145/3319921.3319945

Tekedere, H., & Göke, H. (2016). Examining the effectiveness of augmented reality applications in education: A meta-analysis. *International Journal of Environmental and Science Education*, *11*(16), 9469–9481. https://eric.ed.gov/?id = EJ1118774

Vate-U-Lan, P. (2011). Augmented reality 3D pop-up children book: Instructional design for hybrid learning [Paper presentation]. 2011 5th IEEE International Conference on E-Learning in Industrial Electronics (pp. 95–100). IEEE. https://doi.org/10.1109/ICELIE.2011.6130033

Vate-U-Lan, P. (2012). An augmented reality 3d pop-up book: The development of a multimedia project for English language teaching [Paper presentation]. Multimedia and Expo (pp. 890–895). IEEE. https://doi.org/10.1109/ICME.2012.79

Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry*, *45*(1), 2–40. https://doi.org/10.1046/j.0021-9630.2003.00305.x

Vernadakis, N., Avgerinos, A., Tsitskari, E., & Zachopoulou, E. (2005). The use of computer assisted instruction in preschool education: Making teaching meaningful. *Early Childhood Education Journal*, *33*(2), 99–104. https://doi.org/10.1007/s10643-005-0026-2

Wu, H. K., Lee, S. W. Y., Chang, H. Y., & Liang, J. C. (2013). Current status, opportunities and challenges of augmented reality in education. *Computers & Education*, *62*, 41–49. https://doi.org/10.1016/j.compedu.2012.10.024

Wu, M. H. (2019). The applications and effects of learning English through augmented reality: A case study of Pokémon Go [Paper presentation]. Computer Assisted Language Learning (pp. 1–35). Routledge. https://doi.org/10.1080/09588221.2019.1642211

Yang, H., & Tate, M. (2009). Where are we at with cloud computing? A descriptive literature review [Paper presentation]. 20th Australasian Conference on Information Systems (pp. 807–819). AIS. https://aisel.aisnet.org/acis2009/26/

Yang, J. (2013). Mobile assisted language learning: Review of the recent applications of emerging mobile technologies. *English Language Teaching*, *6*(7), 19–25. http://dx.doi.org/10.5539/elt.v6n7p19

Yang, X., Kuo, L., Ji, X., & McTigue, E. (2018). A critical examination of the relationship among research, theory, and practice: Technology and reading instruction. *Computers & Education*, *125*, 62–73. https://doi.org/10.1016/j.compedu.2018.03.009

Zainuddin, N., & Idrus, R. M. (2016). The use of augmented reality enhanced flashcards for Arabic vocabulary acquisition [Paper presentation]. 2016 13th Learning and Technology Conference (pp. 1–5). IEEE. https://doi.org/10.1109/LT.2016.7562857

Zhenming, B., Mayu, U., Mamoru, E., & Tatami, Y. (2017). *Development of an English words learning system utilizes 3D markers with augmented reality technology* [Paper presentation]. IEEE 6th Global Conference on Consumer Electronics (pp. 1–2). IEEE. https://doi.org/10.1109/GCCE.2017.8229353

Zucker, T. A., Moody, A. K., & McKenna, M. C. (2009). The effects of electronic books on pre-kindergarten-to-grade 5 students' literacy and language outcomes: A research synthesis. *Journal of Educational Computing Research*, *40*(1), 47–87. https://doi.org/10.2190/EC.40.1.c

Author Biographies

Min Fan is an assistant professor in the School of Animation and Digital Arts in the Communication University of China. Her research focuses on innovative technologies to support at-risk children's learning and development.

Alissa N. Antle is a professor in the School of Interactive Arts and Technology at Simon Fraser University. As an innovator and scholar, her research investigates how new forms of interactive technologies can support, change and augment the ways that children learn, develop, and connect.

Jillian L. Warren is a PhD candidate in the School of Interactive Arts and Technology at Simon Fraser University. Her research focuses on technologies to support children's health and well-being, with specific focus on child populations who are sick or chronically ill.