sustainability

MDPI

Article

The Conceptualisation of User-App Interactivity in Augmented Reality-Mediated Learning: Implications for Literacy Education

Tingjia Wang [1,*] and Lynde Tan [2]

[1] Writing Center, Hiroshima University, Higashihiroshima 739-8512, Japan
[2] School of Education, Western Sydney University, Penrith, NSW 2751, Australia; lynde.tan@westernsydney.edu.au
* Correspondence: twang@hiroshima-u.ac.jp

Abstract: Augmented reality (AR) is transforming users' multisensory experiences and heightening the level of engagement with multimodal learning. Scholarly attention is urgently needed to conceptualise and examine user–app interactivity in educational contexts. Drawing on the systemic functional–multimodal discourse analysis approach, this article aims to explore key user roles prompted by AR apps and examine educational functions that these user roles fulfil in AR-mediated learning. Based on our analysis of 14 AR apps selected for a 3-day workshop with six Australian primary school teachers, we identified four categories of user roles that facilitated different literacy activity types during AR-mediated learning. To design effective learning experiences, this article argues that teachers need to consider the resonance between students' AR experiences and their prior engagement with other forms of digital texts when planning for scaffolding strategies.

Keywords: augmented reality; multimodal; user-app interactivity; user roles; scaffolding

check for updates

Citation: Wang, T.; Tan, L. The Conceptualisation of User-App Interactivity in Augmented Reality-Mediated Learning: Implications for Literacy Education. *Sustainability* **2023**, *15*, 10949. https://doi.org/10.3390/su151410949

Academic Editor: Rosabel Roig-Vila

Received: 30 May 2023
Revised: 27 June 2023
Accepted: 29 June 2023
Published: 12 July 2023

1. Introduction

The sustainability of emergent technologies, such as Augmented Reality (AR), requires educational interventions that aim to advance the United Nations' Sustainability Development, Goal 4; this goal emphasizes the importance of education as the enabler of both improvement in people's lives and sustainable development. AR in education contributes to technological innovation, though it can only contribute to quality education when teachers' knowledge in AR-mediated learning is deepened [1]. AR is an emerging and interactive medium that provides an enhanced perception of the real world by overlaying digital 3D visual objects, sound, and other sensory stimuli in the physical environment in real time [2,3]. AR is often accessed via mobile devices, such as smartphones, tablets, and laptops [4,5]. It is widely acknowledged that AR lies on a reality–virtuality continuum [6,7], and three main types are currently available: marker-based, markerless, and location-based AR [8]. Marker-based AR uses a marker, such as visual patterns, a printed QR code, or a sign, to initiate digital animations or augmented experiences when the marker is scanned; in markerless or location-based AR, the AR content is dependent on the user's location, which, in turn, relies on the Global Positioning System (GPS) [9].

AR technology is widely used in gaming (e.g., Pokemon Go), digital marketing (e.g., IKEA Place for placing furniture in one's physical home environment), social media (e.g., photograph filter apps), and many other domains of society. It is increasingly transforming education and creating the need for educators and researchers to investigate its potential benefits for teaching and learning. Recent reviews of AR in education are beginning to report an increased use of AR in K-12 education [10–13]. Key educational benefits include vocabulary learning [14–16], media literacy [17], reading comprehension [18], motivation [19,20], teaching the English alphabet [21], and storytelling [22]. More recently, AR has been shown to develop primary school students' multimodal literacies [23] and agency [24].

AR-mediated learning refers to the teaching and learning practice in which knowledge is conveyed through AR technology [25]. In our research, AR-mediated learning refers to the multimodal literacy practice in which students can develop multimodal literacies through interaction with selected AR stories. Multimodal literacies refer to the capacity to and practice of using two or more modes of meaning in communication [26]. For Southgate et al. [27], AR-mediated learning requires students to interact with virtual objects and characters through gesture, voice, and/or moving around a particular space. When AR is used, embodied meanings are created through gestures, bodily movements, and other modalities [24]. It is through the "naturalised uses of modes" [28] (p. 5) that we can create interactivity in AR-mediated learning. This method requires teachers to pay more attention to the use of space and bodily movement in literacy practices and effectively scaffold students to respond to a wide range of multimodal cues in AR apps.

Embracing AR technology in multimodal literacy practices necessitates the development of teachers' multimodal literacies before they can expand the scope of digital learning and influence students' development of multimodal literacies [23]. However, apart from enriching students' multisensory experiences [29,30] and making literacy practice more interesting, more scholarly attention is needed to assess the effectiveness of AR-mediated learning in relation to the achievement of the following learning goal—how does the selected AR app assist in the achievement of the learning goal in literacy practice? Furthermore, school teachers are also in need of accessible tools to guide their mindful selection and evaluation of AR apps to enhance students' multimodal literacies in AR-mediated learning. To address this gap, this article aims to examine the AR text and the AR experience with a focus on the conceptualisation of user–app interactivity in the educational context. Specifically, this study will be guided by following research questions:

1. What user roles can be prompted via AR instruction for user–app interactivity in AR apps?
2. How can we inform AR-mediated learning through the conceptualisation of user–app interactivity in the educational context?

While existing descriptions of AR apps focus on the semiotic deconstruction of AR texts [31,32], the conceptualisation of user–app interactivity in this study emphasises the role that AR users can play as the text's co-constructor in the interactions between AR objects and characters. This approach will provide a significant perspective to complement the current exploration and understanding of AR as the semiotic text. Our conceptualisation will also provide teachers with a perspective of how to select AR apps. More importantly, it can effectively assist teachers' achievement of educational goals in class and provide them with a trained eye to observe and examine students' user–app interactivity in AR-mediated learning, before providing necessary scaffolding.

2. User-App Interactivity

Previous conceptual descriptions of technological texts include the semiotic conceptualisation of VR typology [30], the development of semiotic meta-language for understanding video game texts [33], and the semiotic description of children's e-books [26]. These frameworks share the semiotic view of AR/VR, i.e., that an AR/VR story is a cohesive text constructed through the interaction between a wide range of linguistic, image, sound, and multisensory semioses.

In this context, two existing conceptual frameworks of AR texts shape the core research context for the present study, and the present study will provide a complementary conceptualisation of the AR text for both existing frameworks. The 3D Manipulation Tasks framework provides a systematic description of the key tasks that users are prompted to complete based on the AR instruction during user–app interactivity [31]. The framework identifies four key tasks: Selection, Positioning, Rotation, and Scaling. Selection refers to identifying a particular virtual object or part of an object; Positioning refers to changing the 3D position of a virtual object or character; Rotation refers to changing the direction of a virtual object; and Scaling refers to changing the size of a virtual object [31]. The 3D

Manipulation Tasks framework categorises the most basic user–app interactivities with a focus on the user's action on the app. The framework, however, does not discuss how these manipulation tasks are prompted in the AR app through what kind of multimodal cues. The investigation of the multimodal cue in the AR instruction is significant to general users' multimodal literacies, as although the AR technology has emphasized the multisensory experience users can obtain in the user–app interactivity, the potential of the user's manipulation of the virtual character or object is strongly based on the prescribed AR instruction. In other words, when and what manipulation tasks that users are allowed to conduct are typically prescribed by the AR story. Although users can conduct the four categories of manipulation tasks on virtual characters and objects, users must follow the AR instruction and provide the prescribed action or verbal input to process the AR experience. Without the effective response to the AR instruction, the AR experience is typically interrupted, and the AR text cannot be unfolded smoothly. Thus, the 3D Manipulation Tasks framework needs another complementary description of the typical AR instruction for prompting the four manipulation tasks.

Another gap of the 3D Manipulation Tasks framework is that it focuses on the user's interactivity realised through fingers and does not describe the interactivity realised through the user's bodily movement, such as walking and tilting the iPad. The description of the user's bodily movement is significant to both the assessment of users' multimodal literacies and the multisensory potential in the AR app. One feature distinguishing the AR experience from other multimodal experiences like using the mobile phone, reading an e-book, or watching a film is that the user–app interactivity in the AR app tends to involve the bodily movement in the real, physical space. By walking in the real space while holding the iPad, the user is navigating in the virtual space. That is, the spatiality and the directionality are two key multimodal literacies that an effective AR user should have; meanwhile, spatiality and directionality are two key multimodal literacies that students can significantly develop through the AR-mediated learning. This further justifies the necessity of involving AR apps in the literacy education.

The other prior conceptual framework related to the user–app interactivity is the description of interactivity and narrative functions based on a systematic investigation of narrative AR texts [32]. Drawing on Social Semiotics, the user's interactivity like tap, pinch, reverse, drag, and swipe are viewed as the user's co-construction of the AR story in this conceptual description [32]. This conceptual description argues that by providing different interactive responses, the user can influence the outcome of the AR story. However, this argument is relied on the design of the AR text. In some AR narratives, the app opens up the meaning-making potential for the user to be the co-creator of the story. Users can experience the AR story several times and try different interactivity choices to explore how their interactivity can construct the AR experience differently in the same app. Similar to the 3D Manipulation Tasks framework, this conceptual description also emphasises the user's interactivity and influence on the AR text. As the other side of the user–app interactivity, the role of the AR instruction is lack of exploration. Some key questions for the description of the AR cue in the user–app interactivity remain un-answered. These questions include: When and through what multimodal instructions does the AR send the invitation to users to co-construct the AR story? How salient or tacit the AR invitation can be? Can the majority of users identify the AR cue and effectively play the co-constructor role? What multimodal literacies are required in the user to understand the AR instruction and to provide the effective response to the AR invitation?

To respond to these questions, a clearer conceptualisation of the user–app interactivity is needed. This conceptualisation should describe both the user's typical interactivity choices and the common multimodal cues in AR apps. To fill the gap in both prior conceptual frameworks, this article aims to explore the role users can play in the user–app interactivity, with a focus on the typical AR instruction instantiated in AR apps for prompting the user's interactivity. The description will enhance the general user's multimodal literacies in understanding multimodal cues in the AR app. This article will also present the emphasis on the AR app, which shows the potential as the teaching material for the AR-mediated literacy education. The conceptualisation of the user–app interactivity proposed in this article will enhance both teachers' and students' multimodal literacies.

3. Materials and Methods

3.1. Participants and Data

This study was conducted as part of a larger-scale project on the use of AR to teach language and literacy in the primary school context in Australia. In this paper, the study was based on our observation and analysis of six Australian primary school teachers' interactions with 14 AR apps selected for a 3-day workshop in 2020. Prior to the workshop, the second author, who was the Chief Investigator of the larger-scale project, searched for AR apps on App Store and Google Play. Given the research focus on primary language and literacy teaching, only apps that offered the potential for this disciplinary focus were chosen for the workshop.

As part of the objectives of the workshop, the teachers were introduced to the AR technology and the reported educational affordances of AR from past educational research. These teachers experimented with a range of free or low-cost AR apps (see Table A1 in the Appendix A) and evaluated their suitability for developing primary school students' multimodal literacies. They also evaluated the relevance of the AR apps to the targeted units of work that they intended to implement for students across different year levels for a school term that year (see [23] for a full report of the teachers' evaluation of these apps which is beyond the scope of this paper). Hence, the workshop was designed to provide these teachers the opportunity to experience AR apps which could be potential digital resources for teaching language and literacy in the Australian context. The workshop also aimed to observe and describe the existing multimodal literacies in teachers based on their interaction with AR apps. After the 3-day workshop, the teachers then selected a few AR apps that they deemed suitable for enhancing the units of work that they planned to redesign and implement.

As noted earlier, multimodal literacies refer to the capacity and practice of using two or more modes of meaning in communication [34]. The multimodal literacies shown in teachers' interaction with AR apps in this workshop are mainly developed through their prior experiences with other forms of multimodal texts, such as picture books, video clips and films. Teachers are familiar with and experienced in selecting and implementing these multimodal texts as teaching materials in their class to assist their achievement of the learning goal. In terms of the AR app as a text type, we found that teachers tend to rely on their prior multimodal experiences with other forms of texts to understand and to respond to the instructions provided in AR apps.

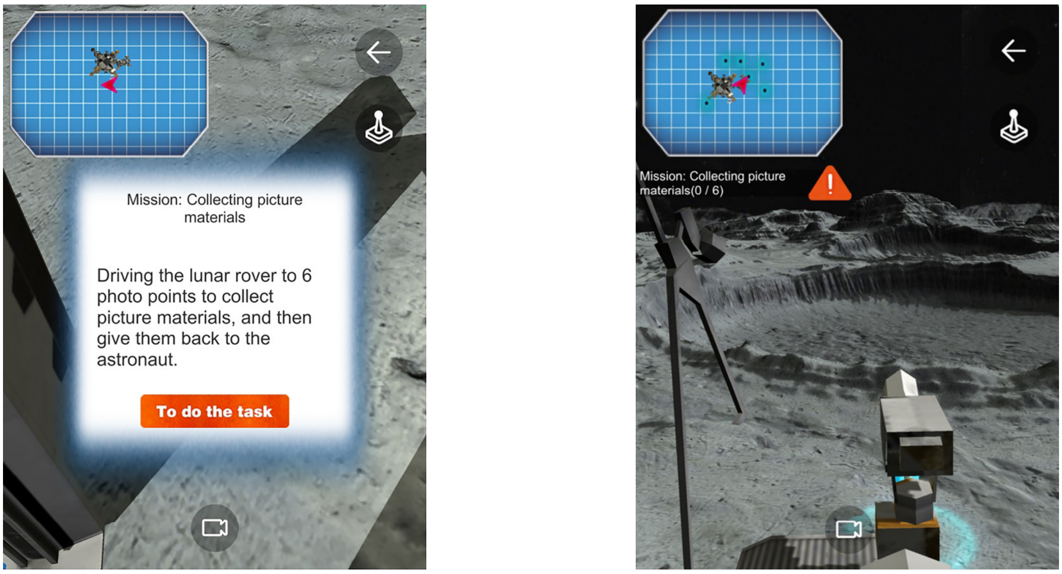
The recruitment of the teachers was based on purposive sampling. They were all from the same school and had past professional relationships with the second author. These teachers were interested in embracing AR in their teaching and learning practice, yet needed more multimodal lenses to guide their selection of AR apps that suited their classes and learning goals.

3.2. Data Analysis

The analyses undertaken were based on the video recordings of teachers' experimentation with the 14 AR apps and screenshots of the AR apps that showed different forms of touch designs. The analytical approach applied to the AR app is known as the systemic functional–multimodal discourse analysis approach [35], drawing on the Social Semiotics theory [36]. The Social Semiotics theory views a cohesive multimodal text as the result of the interaction across selected multimodal semioses in the text, such as the language, the image, the music, and the editing [36]. In other words, a cohesive multimodal text is more than the simple sum of multimodal semioses; rather, it is the integration of these semioses that makes the text cohesive and meaningful [37]. As the semiotic technology, an AR text is the result of the meaningful interaction among a number of semiosis, such as the language, the image, the layout, and the movement. Conceptualising the AR text as a semiotic text can therefore foreground the meaning-making nature of the technology and assist teachers' evaluation of the AR text in relation to the learning goal and the literacy development [38–41].

One strength of the systemic functional–multimodal discourse analysis approach is that it integrates two discourse analytical approaches (i.e., the systemic functional approach and the multimodal discourse analysis approach) and allows researchers to conduct a balanced analysis of the meaning-making choice made in the language modality and the non-language modality, respectively. The balanced analytical focus is particularly suitable for the analysis of multimodal texts and to explain how the cohesive meaning is constructed in a multimodal text, such as an AR story, through the interaction among a wide range of audio-visual resources such as language, image, sound, and touch. Specifically, the systemic functional–multimodal discourse analysis approach includes two discourse analytical approaches. The systemic functional approach refers to the linguistic theory and analytical approach known as Systemic Functional Linguistics [42–44]. The analytical tool in Systemic Functional Linguistics is for the description of the meaning constructed in the modality of language [45]. The multimodal discourse analysis approach refers to the analytical approach for the description of the meaning constructed in non-linguistic modalities, such as image, sound, editing, and camera movement [46–48]. Bringing the systemic functional approach and the multimodal discourse analysis approach together, we can analyse the cohesive meaning constructed, not only in each of linguistic and non-linguistic modalities, but also the meaning constructed through the interaction across modalities [49,50].

The systemic functional–multimodal discourse analysis approach is exemplified by the sample analysis of AR Moon, an AR app used in our 3-day workshop (see Figure 1). The analysed scene is the first mission in the AR Moon, illustrated by two key moments in Figure 1a,b. Figure 1a is the AR instruction provided to users to introduce the mission and to start users' AR experience on the virtual Moon as an astronaut. Figure 1b is the user's perspective driving the virtual lunar rover while following the instruction in Figure 1a and completing the mission.



(**a**) (**b**)

Figure 1. AR Moon first mission scene. (**a**) The AR instruction; (**b**) The user's perspective driving the virtual lunar rover on the virtual Moon while completing the mission described in image (**a**).

As shown in Figure 1, to effectively respond to the AR instruction, the user needs the strong multimodal literacies about AR as a semiotic text in four key aspects: the language, the image, the bodily movement and the directionality. The user needs to link the linguistic property name "lunar rover" in the instruction with the image of the virtual lunar rover during the mission, and the user needs to understand that the linguistic activity "driving" in the instruction will be realised by the user's walking while holding the iPad in the bodily movement. One essential multimodal literacy commonly requested in users in the AR experience is the capacity of using the bodily movement to respond to the AR instruction. The user's spatiality in the real, physical world and the user's capacity of translating the AR instruction into real actions are significant to the effective user–app interactivity in the AR experience.

While "driving" the virtual lunar rover, the user then needs to find the six pictures on the virtual Moon landscape by following the map on top (see Figure 1b). The pink arrow indicates the current location of the user and the rover that the user is driving; the six black dots refer to the six pictures that the user needs to collect, i.e., the "6 photo points" in the AR instruction; the large brown shape on the map refers to a virtual rocket where a virtual astronaut is waiting for the user to bring back the six pictures. This virtual astronaut is the "astronaut" in the AR instruction in Figure 1a. The directionality, another essential multimodal literacy, is then prompted in this user–app interactivity. The user is supposed to be able to read the map and use the map to guide the bodily movement, i.e., walking while holding the iPad, to reach to the six photo points. After visiting the six photo points, the user is supposed to follow the map and walk back to the point of departure, i.e., the virtual rocket spot on the map. The user's walking to the six photo points is considered as "collect picture materials" in the virtual space, and the user's walking back to the departure point behaviour is considered as "give them back to the astronaut" in the virtual space. The capacity of translating the linguistic instruction to the bodily movement is again requested as one of the essential multimodal literacies in the user–app interactivity. Other necessary multimodal literacies requested in this example include the user's directionality and spatiality. Directionality and spatiality are not commonly requested in the user's interactivity with other forms of multimodal texts because users can simply consume the multimodal text like a film or a picture book while sitting. However, while consuming AR texts, the user's directionality and spatiality are essential multimodal literacies to the effective user–app interactivity.

Using the systemic functional–multimodal discourse analysis approach, we can identify the key linguistic and image semioses in this sample AR text and the interaction among these semioses. The key item "lunar rover" has been triple-coded through the language, the image, and the symbolic arrow on the map. Another key item "picture materials" has been double-coded through the language and the symbolic black dots on the map. We can also analyse teachers' bodily responses to the linguistic semiosis to assess teachers' existing multimodal literacies. This assessment result can provide us a clue to develop metalanguage tools for teachers to assist their development of AR-specific multimodal literacies and assist teachers to facilitate students' multimodal literacies in AR-mediated learning. Taking the systemic functional–multimodal discourse analysis approach, we report the four key categories of user roles prompted in the user–app interactivity in AR apps and suggest the potential educational function of each user role in the AR-mediated learning in the rest of this article.

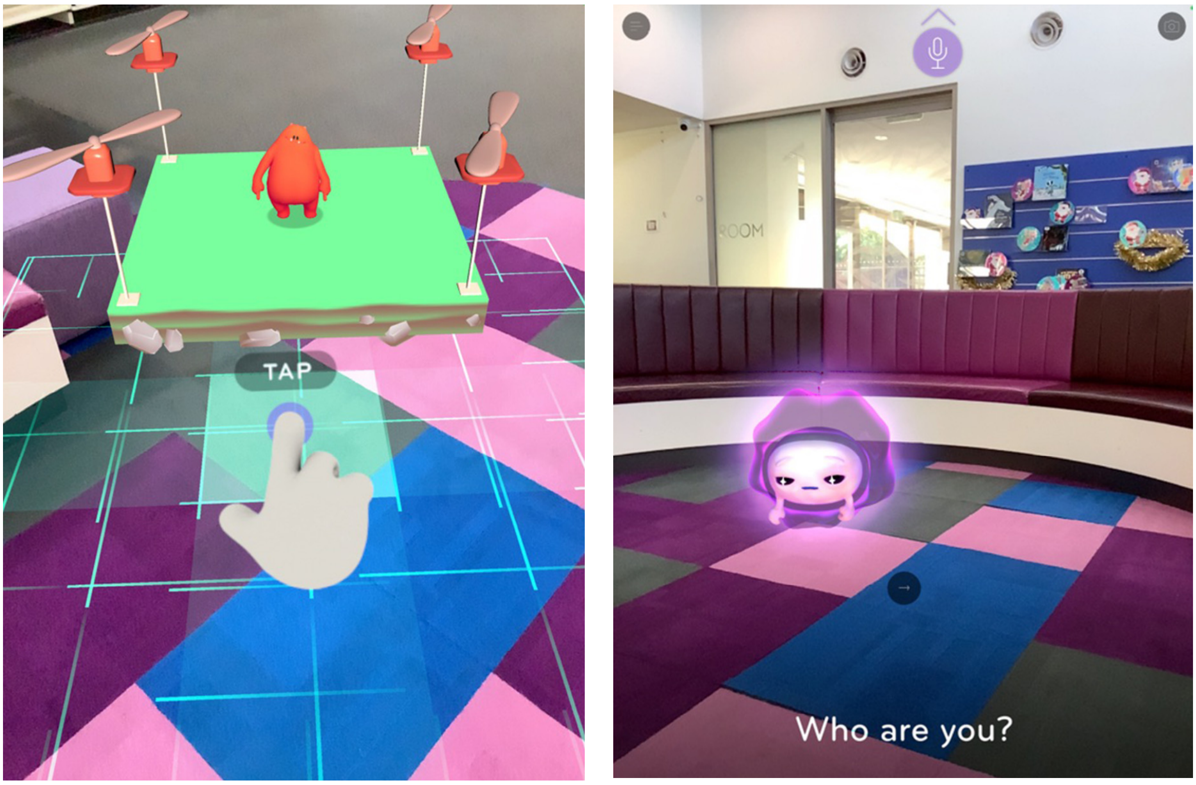
4. Results

The four key user roles prompted in the AR app in our analysis include: the trigger, the viewer, the manipulator and the creator. As noted earlier, users tend to rely on their prior multimodal experiences to respond to the user–app interactivity in the AR experience, especially when the user is new to the AR technology or new to the AR app. Additionally, materiality of the device and the technology used in the educational context can affect the learning practice [34,51]. In this section, we will describe each user role with examples, and discuss the typical prior multimodal experience users tend to rely on during the AR experience. Our description and discussion will show a wide range of audio-visual semioses used in AR apps to invite users for the user–app interactivity. Our discussion will highlight AR-specific multimodal literacies that are needed for the user to be an effective AR user.

4.1. The Trigger

In most AR experiences, the first role that the user is invited to play is the Trigger. The Trigger role is typically for starting the unfolding of an AR text, as well as the user's own AR experience. The interactive nature of AR makes the AR app often provide users the opportunity to co-construct the AR experience with the app, and the multisensory nature of AR makes the AR invitation for users' triggering responses diverse. The AR invitation for users' triggering responses can be sent through image and/or language, and the channel for users to respond can be the bodily movement and/or the verbal input.

Figure 2 presents two examples of AR invitations and users' responding channels in the Wonderscope app. The two invitations are sent to users for the triggering response at different phases of the AR story Clio's Cosmic Quest. As the AR story is designed to be unfolded phase by phase, the AR experience is also shaped as different phases, and to start each phase of the AR experience, the user is invited to play the Tigger role within the app. Figure 2a is a typical invitation for users to trigger the AR text by tapping on a certain button on the screen; Figure 2b is a typical invitation for users to trigger the AR text by providing a prescribed verbal input.



(**a**) (**b**)

Figure 2. Examples of AR invitations for the user role "Trigger". (**a**) Invitation for users' bodily response in the Wonderscope; (**b**) Invitation for users' verbal input in the Wonderscope.

As shown in Figure 2a, to prompt the user's role as the Trigger, multimodal semioses integrated in this AR invitation include the language "TAP", the image (i.e., the button), and the illustration of bodily movement (i.e., the tapping hand). The most salient instruction provided in this invitation is the tapping hand. It is not simply a virtual image; it is also a translation of the linguistic instruction "TAP" to the user's bodily movement. Compared to the linguistic instruction "driving the lunar rover" in Figure 1, this tapping hand makes the AR invitation more beginner friendly. In our observation of teachers' responses, teachers can easily respond to this invitation and tap the screen to start the AR text. Although tapping on any part of the screen can trigger this AR text, some teachers in our observation tend to tap on the exact point as the virtual hand indicates. Another reason for teachers' quick, effective responses to this TAP invitation is that the requested bodily movement, i.e., tap on the screen, is an everyday behaviour in this digital era. Users have been tapping on their mobile phones and other digital devices every day today. Prior to the TAP invitation, as part of the instruction to start the AR story, the user is supposed to find a flat surface with sufficient light to project virtual characters. The grid in Figure 2a is the sign that the AR app is assessing the surface and the lighting condition that the user selected for triggering the AR text, and as the surface is flat and the lighting condition is sufficient, virtual characters can then be projected into the physical, real context.

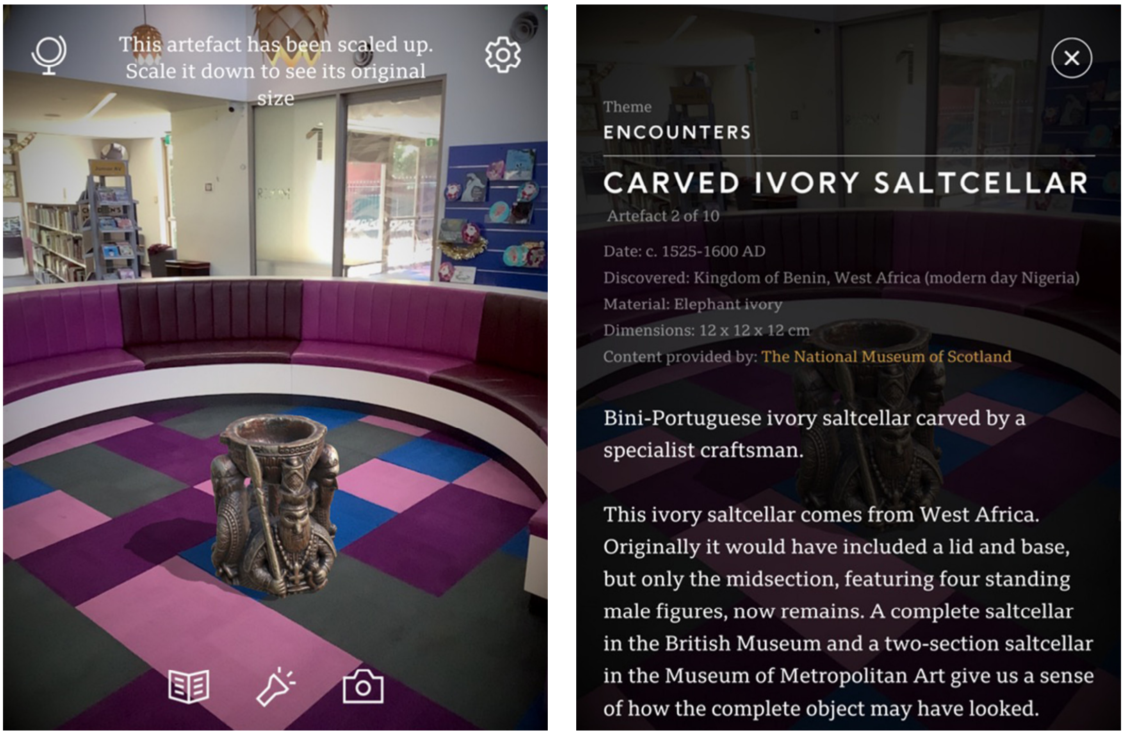
Figure 2b illustrates another invitation sent to users in the same AR story. The triggering invitation was sent through the language "Who are you?" at the bottom of the app. The user is supposed to press the microphone button on top of the app and read the displayed question "Who are you?" loudly to the device. The main multimodal semiosis integrated in this invitation include the language and the microphone icon. This invitation is designed to trigger the user's interactivity with the purple, virtual character, Clio, in the AR story, as shown in Figure 2b. The question "Who are you?" is designed to be asked by the user to Clio to initiate the user–character interactivity. In other words, the app invites the user to talk to the virtual character first, which is a strategy to enhance the user's sense of engagement in the story. To effectively respond to this invitation, the user should have the capacity of identifying the microphone button on top and understanding that the prescribed verbal input is the question: "Who are you?". If the user misread the invitation and thought that the question was asked from Clio to the user, the user may input his or her own name as the verbal response. Without verbal input or with the verbal input other than the exact question, "Who are you?", the story cannot be triggered. In our observation, teachers can effectively respond to this invitation.

The Trigger role is typically played at the beginning of the AR text or the beginning of a phase of the AR text. Touch is the effective practice for initiating the body–screen relationship in this digital communication [52–55]. Likewise, the voice input also plays the role of initiating the body–app relationship as the common triggering form in the AR text. The AR initiation for this user role tends to be beginner friendly and not request a high level of multimodal literacies in its users. The requested responses from the user to trigger the app are typically prescribed, yet simple, responses, which are similar to users' everyday interactivity with other digital devices and can be accomplished easily. Inviting users to play the Trigger role in the AR experience can effectively increase users' sense of engagement with the AR story and, more importantly, can familiarise users with the multisensory interactivity that the AR app requests. Requests like finding a flat surface with sufficient lighting and providing the verbal input can let users realise that, in addition to finger actions like tapping and swiping, the bodily movement and the verbal interaction are common behaviours involved in the AR experience, and that spatiality about the physical, real environment is needed as one of the AR-specific multimodal literacies.

4.2. The Viewer

The second category of the user role prompted in the AR app is the Viewer, i.e., the user is positioned at the observing role to receive messages provided by the AR app. The Viewer role typically occurs in AR apps that show strong informative, educational purposes, or when the user needs the information as the preparation for the follow-up user–app interactivity. For instance, the user needs to watch the unfolding of the AR story before the user–character interactivity. Playing the Viewer role in AR apps is similar to the experience of reading an e-book or watching a film. All of these situations place the user in the position of the information receiver and, subsequently, the app becomes the information provider.

Figure 3 illustrates the scene where the user is positioned at the Viewer role in the AR app called BBC Civilisation AR. It is an educational app for users to learn knowledge about historical artefacts worldwide. The app can be viewed as a digital, interactive archive of historical artefacts in the world or, more precisely, a digital, interactive museum. Figure 3a is an example of the archived historical artefacts users can see in this app, as exemplified by the carved ivory saltcellar. There are three icons, i.e., the book, the flashlight, and the camera, at the bottom of Figure 3a. By tapping the book icon, users can read the passage about the saltcellar, as shown in Figure 3b.



(**a**) (**b**)

Figure 3. Example of the user role "Viewer". (**a**) The virtual carved ivory saltcellar in BBC Civilisation AR; (**b**) Information displayed about the carved ivory saltcellar after tapping the book icon at the bottom of image (**a**).

As shown in Figure 3b, the information about the carved ivory saltcellar is displayed by the user tapping the book icon in Figure 3a. After reading the passage, the user can simply tap the cross button on top to close the information page and head back to the previous interface (Figure 3a). The user can also tap other buttons at the bottom in Figure 3a. The flashlight button can provide a piece of verbal passage about the saltcellar, which is a simplified version of the passage in Figure 3b. By tapping the camera button, the user can take a photo of the virtual artefact in the physical context and the photo will be automatically saved in the device. All these built-in functions serve the same purpose of making users' AR experiences closer to a real museum visit. The experience of reading the displayed passage in the app is the same with the experience of reading the passage displayed next to an artefact in a real museum. Many museums also provide audio versions of the passage for visitors to listen to and allow visitors to take photos of the artefact.

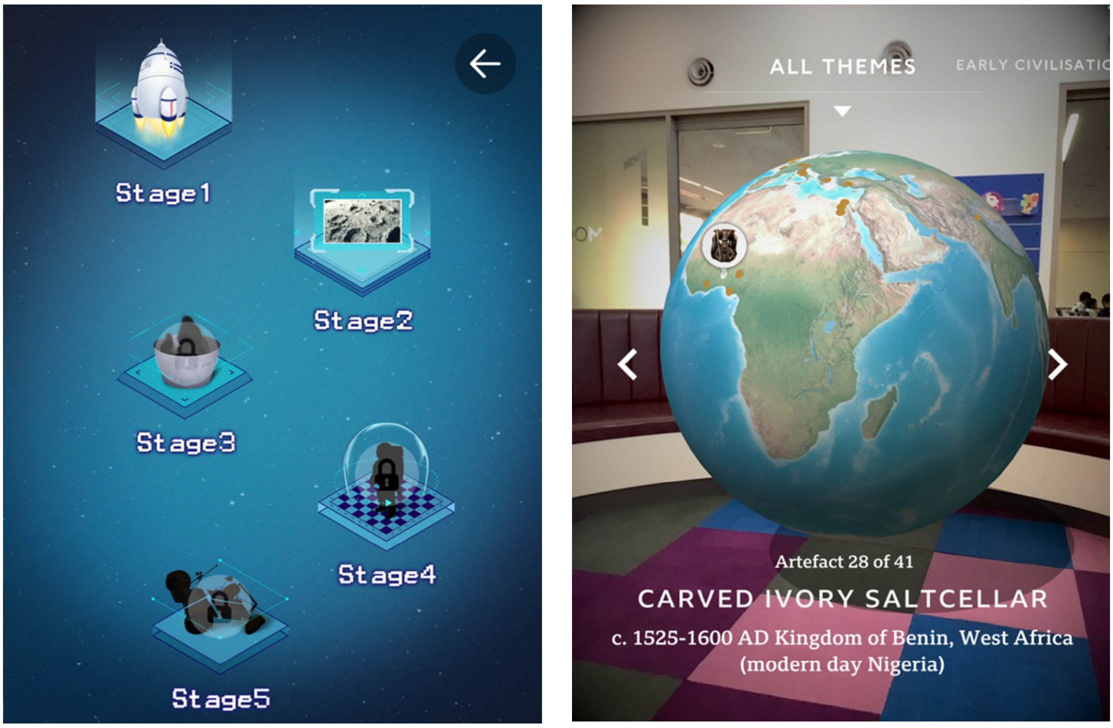
This app has other designs to resonate users' AR experiences with their prior experiences of visiting a real museum. One example is the shadow of the saltcellar in Figure 3a. The shadow is part of the design of the virtual saltcellar in the AR app and is not affected by the lighting in the physical context. The shadow increases the authenticity of the virtual artefact. Another example is that users can use fingers to pinch on the screen to zoom in and see the detail of the artefact. This pinch movement is for the similar purpose as leaning forward to see the detail of the artefact at a real museum. Likewise, the user can change the iPad angle to see different sides of this 3D saltcellar. Compared to 2D pictures of artefacts displayed on the museum website, this AR app resonates better with our prior real-world experiences in a museum.

Although playing the Viewer role involves less proactivity from the user compared to other user roles, the viewing process still requires physical engagement in the interaction with the AR app, as the user needs to keep holding the screen at a certain angle and sometimes also needs to rotate the body to view different sides of the virtual object [56]. This distinguishes the viewing experience in AR apps from the viewing experience of other digital texts like films or TV dramas. In our observation, teachers can easily play the Viewer role as this user role is typically not a choice by the user but prescribed by the app. Another reason for teachers to be comfortably positioned at the Viewer role in the AR experience is that it is a common role people play in everyday life while absorbing information. While playing the Viewer role, the AR experience shows the minimum multisensory interaction.

4.3. The Manipulator

The third category of the user role is the Manipulator, which means that the user can construct their own unique AR experiences during the user–app interactivity, and that different users can construct different AR experiences in the same app. As noted earlier, in the previous descriptive framework of AR narratives, user's manipulation function has been identified and described, yet with a focus on the user's influence on the outcome of an AR narrative [32]. In our conceptualisation, the Manipulator role refers to the user's influence on their own AR experiences. One condition for users to construct their customised AR experiences is that the AR text cannot be designed and unfolded in a linear format. For an AR text designed in a linear format, the user has to follow the prescribed unfolding order of the AR story and cannot choose the departure point in the story. To allow users to customise their AR experiences, the AR text needs to be designed in the phased form or the thematic style. The AR text also should provide the menu scene for users to select the departure of their AR experiences and provide users the opportunity to determine the order of phases or themes they want to experience in the AR text. In such an AR experience, the user is not simply a receiver or a consumer of the AR text but also a co-constructor of the AR story. This is a higher level manipulation of the AR text, which is beyond the simple re-sizing or rotation of the virtual character or object.

Figure 4 illustrates two examples of the AR design, providing users with the opportunity to play the Manipulator role. Figure 4a is the menu scene of the AR Moon app; Figure 4b is the menu scene of the BBC Civilisation AR app.



(**a**) (**b**)

Figure 4. Examples of the menu scene for the user role "Manipulator". (**a**) The menu scene of AR Moon; (**b**) The menu scene of BBC Civilisation AR.

Figure 4a,b are examples of the phased AR experience and the thematic AR experience, respectively. As shown in Figure 4a, the exploration of the virtual Moon in this app is phased into five stages and users can choose the departure stage as well as the order of these five stages, although the built-in order of the stage from Stage 1 to Stage 5 indicates that there is a recommended version of the AR experience. Users can also choose not to customise their AR experiences and simply follow the recommended order from Stage 1 to Stage 5. The illustration in Figure 1 is selected from Stage 2 in this app. The AR experience of customising the order of the five phase in the virtual Moon expenditure is closer to the experience of reading a book without following the chapter order or the experience of watching a television drama without following the episode order. In our observation, although the meaning potential of customising the AR experience is opened, teachers tended to follow the Stage order in the app from Stage 1 to 5. One reason is that teachers were new users to this app and preferred this approach to accepting the recommended Stage order, which is also the recommended version of the AR experience. Another reason is that the Stage order is the salient visual message in the menu (see Figure 4a). The Stage order is double-coded in the modality of language by the number 1 to 5 and the modality of image by the top-down display. This multimodal presentation of the menu provides users the impression that there is a prescribed order of the five stages and that their AR experiences are also prescribed. Some users follow the Stage order due to the unawareness that they are indeed allowed to customise their AR experiences in this app. Similar visual presentation styles can be seen in Table of Contents of a book or the episode menu of a drama.

The menu scene of the BBC Civilisation AR app, on the other hand, illustrates the thematic presentation style without any recommended order by the app (see Figure 4b). Users can choose the theme on top to determine the departure of their AR experiences. As shown in Figure 4b, the carved ivory saltcellar illustrated in Figure 3 is an artefact selected under the theme "EARLY CIVILISATION". Within each theme, users can further choose the departure of their AR experiences based on the continent or the exact artefact they want to see at first. For instance, the saltcellar is the artefact No. 28 from West Africa. Users can tap the leftward and the rightward arrows to rotate the virtual globe to see other artefacts and continents under the selected theme. The AR experience in this app is totally customised, and the AR experience is even more customised than the real visit of a museum. Both the AR experience and the visit of a real museum allow the customer to choose which artefact to see at first and the order of all artefacts the customer wants to see. However, during the visit of a real museum, the visitor's tour is typically influenced by the space setting, such as the level of floors and the display arrangement of artefacts. Visitors tend to visit artefacts from the ground floor and from the artefact displayed near the gate. The AR app excludes these spatial factors and allows users to freely choose artefacts they want to visit in a customised order. The potential of customising the users' visiting experience to artefacts has been opened up to the maximum extent in the BBC Civilisation AR.

In our observation, teachers tended to explore the theme list on top at first and then rotate the virtual globe to see all artefacts under the selected theme before deciding on the first artefact to visit; teachers started their AR experience in this app with different artefacts. The Manipulator role afforded by the AR technology plays a significant role in changing the literacy education. On the one hand, it involves learners into the co-construction of both the targeting knowledge and the learning experience; on the other hand, it provides users the unique experience of using touch and fingers as the main exploring method in the literacy learning [57–59].

4.4. The Creator

The prior three categories of the user role, i.e., the Trigger, the Viewer and the Manipulator, focus on the user–app interactivity in AR apps, positioning users as the consumer of the AR text. AR apps in this trend need to provide the built-in content for users to interact with. Shifting the focus from the AR text consumption to the AR text composition, we found the fourth category of the user role, the Creator. The Creator refers to the role users play in a particular type of AR apps which do not provide prescribed stories, but serve as an AR platform for users to create their own content. Apps like these usually provide built-in libraries with virtual character or object images and sound effects, and allow users to import their self-made videos or photos.

Figure 5 illustrates an example of making the primary school office into a dental clinic in the Thyng app. In our 3-day workshop, teachers only experienced the built-in libraries in Thyng and did not have opportunities to create their own AR content. However, teachers have implemented this AR app in their literacy class after the workshop and students have created their own AR texts using the Thyng app. In the AR-mediated learning, students were asked to re-imagine the school space and create a short AR text. Figure 5 is the illustration of a student's reimagination of the school office as a dental clinic. Figure 5a is the photo of the school office; Figure 5b is the illustration of the student's AR text created in the Thyng app.

As shown in Figure 5b, to reimagine the space, the student overlayed the television screen with a photo of a dentist and a child patient. The student asked a peer to sit and act as a child patient waiting for the doctor in the dental clinic. To make the AR text, the student needs to video record the peer siting in the office scene, and then import the video into the Thyng app to overlay the screen with the photo. The student also drafted a short script and audio record the script. The student then added the audio record to the video in the Thyng app as the voice over.

(**a**) (**b**)

Figure 5. Example of the user role "Creator". (**a**) The photo of the school office; (**b**) The AR text making the school office into a dental clinic in the Thyng.

Playing the Creator role foregrounds the developing direction in the literacy education, as learners become content creators rather than passive viewers. Moreover, content making apps like Thyng provide the literacy education learning spaces for the content-making-mediated representation and interpretation [57,60,61]. To play the Creator role, users need a high level of AR-specific multimodal literacies to mindfully select and integrate audio-visual semiosis, in order to construct a cohesive AR text. As a new form of "reality", AR overlays the virtual reality on the real, physical context [62,63]. Users thus need a high level of multimodal literacies to view, represent and interpret the layering of virtual and real "reality" as a new semiotic modality in AR, which is uncommon in other forms of multimodal texts. The closest multimodal experience students can rely on during the play of the Creator role and the composition of their AR texts is the experience of making the video content, such as vlogs. However, the vlog making experience cannot fully support students' use of AR content making apps. One significant feature distinguishing the AR text from other audio–visual texts is the layering between the virtual character or object and the real context. Among the four user roles prompted by the AR app, the Creator role is the most distant role from users' prior experiences with other forms of multimodal texts, which explains why teachers felt that the composition of their own AR content is the most challenging task in the workshop.

5. Discussion

To enrich the students' learning experience and enhance their multimodal literacies, teachers have to be trained to introduce the AR innovations into their teaching design [64–66]. The four categories of the user role provide teachers a framework to understand the user–app interactivity in the co-construction of the AR experience. As we observed, teachers tended to rely on their prior experiences with other multimodal texts to understand AR instructions and to interact with AR apps. Linking the targeting knowledge to students' prior knowledge is a common scaffolding strategy in the literacy practice. Thus, the first suggestion to teachers in the design and the implementation of

AR-mediated learning is to assess students' prior multimodal knowledge obtained through their interactions with other forms of multimodal texts. Teachers are suggested to guide students to demonstrate and reflect their prior multimodal literacies with other forms of digital texts and to link these prior multimodal literacies to the viewing and representing literacies that are needed in the interactivity with the AR text [67]. With the awareness of the four categories of the user role and the resonance between each user role and the prior multimodal experience, the teacher should be able to assess if the student needs any preparation about the user–app interactivity prior to AR-mediated learning. The teacher should also have the capacity of monitoring students' user–app interactivity during AR-mediated learning, and be able to realise the reason behind students' ineffective responses to AR instructions [68].

The common reason for the failure in responding to AR instructions is that the requested multimodal literacy in the user–app interactivity is too distant from the user's prior multimodal experiences. The distance may be due to the fact that the AR technology has the affordance that other forms of technology do not. For instance, the AR composition requires users to have the strong multimodal literacies about the layering of the virtual and real objects and environments, which is the category of multimodal literacies that users can rarely develop in their prior interaction with other forms of multimodal texts. Another reason for the distance between the AR experience and the prior multimodal experience can be the users' limited engagement with digital texts previously, and if the user has not developed sufficient multimodal literacies to support the smooth engagement with AR text. We suggest that, as preparation for AR-mediated learning, teachers should negotiate the glossary with students and learn students' shared prior multimodal literacies. During the AR-mediated learning, students' ineffective user–app interactivity is typically signalled by the interruption in AR text, such as if the AR story stops unfolding due to the users' incorrect input, or a new phase of the AR story is unexpectedly triggered due to the users' incorrect input. We suggest teachers to monitor students' interaction with the AR app, with the awareness of the existing resonance between the required AR-specific multimodal literacies and students' prior multimodal literacies.

In addition to taking students' prior multimodal experiences into consideration, another suggestion to teachers is that they should focus on the mindful selection of AR apps that suit the learning goal. The involvement of AR in literacy practice should service the purpose of further developing students' multimodal literacies through this emerging technology, rather than simply making the class more interesting and digital. One topic that is commonly ignored and has a lack of sufficient discussion in AR-mediated learning is the educational function of the AR app implemented in the literacy practice: How does the selected AR app assist the achievement of the learning goal? To address this question, teachers need meta-awareness about the educational function realised in the user–app interactivity. Drawing on the speech function theory in Systemic Functional Linguistics [45], this article proposes a systematic description of the educational function of the four user roles as a conceptualised framework for teachers to assess the effectiveness of the selected AR app, in relation to the learning goal.

The speech function in Systemic Functional Linguistics was developed based on the description of the verbal exchange, with a focus on the speech role that each speaker plays in the dialogue [45]. The speech function theory is premised on Berry's concepts of the primary and secondary knowers and actors [69,70]. The primary knower (K1) refers to the speaker who already knows the information; the primary actor (A1) refers to the speaker who will actually conduct the action [45,69,70]. In other words, the primary knower (K1) is the source of the information in a dialogue, and the secondary knower (K2) is the information receiver. The primary actor (A1) is the doer of the action commanded in a dialogue, and the secondary actor (A2) is typically the person commanding the action. For instance, when the teacher asks a student, "Can you share the iPad with your peer?", the teacher is A2, and as A1, the student can either respond by the language "Yes, I will.", by the action (i.e., directly share the iPad with the peer) or both (i.e., verbally respond "Yes, I will" and share the iPad with the peer at the same time).

The speech function framework was originally developed for the description of the exchange in the speech role among K1, K2, A1 and A2 in a dialogue, which provides linguists a perspective to interpret the ownership and the exchange of the power in a conversation. Drawing on the four speech functions, we describe the educational function of the four user roles, as follows. When the user plays the role as the Trigger, the user is the primary actor and the secondary knower (A1 and K2), as the user is the actual doer of the trigging action; yet in terms of the AR text, the user is the receiver of the information. When the user is at the Viewer position, the user is the secondary knower only (K2), as the AR app does not invite the user to provide any behavioural input, and the user is positioned at the observing role as the information receiver. When the user is the Manipulator, the user is the primary actor and the secondary knower (A1 and K2). On one hand, the AR app opens the meaning-making potential for the user to customise the AR experience, and the user is the actual doer in customising the AR experience by selecting the departure point in the AR text and the order of the phases/themes in the AR experience. On the other hand, the user is still the information receiver because the AR story is prescribed in the AR app. When the user plays the role as the Creator, the user is the primary actor and primary knower (A1 and K1) since the user is the actual doer of composing the AR text and the user is also the owner of the information that will be composed into the AR text.

The four user roles, the prior multimodal experience resonating with each user role, and the conceptualisation of the educational function of the four user roles, have been presented in Table 1.

Table 1. Conceptualisation of the user role in AR-mediated learning.

User Role	Prior Multimodal Experience	Educational Function
The Trigger	the tap or verbal input in a mobile phone	Primary Actor & Secondary Knower (A1 & K2)
The Viewer	the passage reading or film watching experience	Secondary Knower (K2)
The Manipulator	the book reading experience, regardless of the chapter order or the drama watching experience without following the episode order	Primary Actor & Secondary Knower (A1 & K2)
The Creator	no prior multimodal experience can be linked to	Primary Actor & Primary Knower (A1 and K1)

As the emerging multisensory technology, AR typically positions the user at the primary actor (A1) role. The positioning can be realised either through the clear instruction to prescribe and invite users' actions or by opening up the meaning-making potential for users to interact with virtual objects and characters on the screen (e.g., the menu example in Figure 4). Roles like the Trigger, the Viewer and the Manipulator are typically prompted in the AR app, providing the prescribed text and the user are typically the consumer of the AR text. In AR apps providing the content making platform, like Thyng, the user is provided the opportunity to become the primary knower and actor at the same time as the user is the AR composer. As noted earlier, in our observation, teachers felt that composing their own AR texts is the most challenging task in the user–app interactivity. One reason provided earlier is that the composition of AR texts requires the AR-specific multimodal literacies about the layering between the virtual and the real contexts, and this AR experience is the most distant from most users' prior multimodal experiences. Another reason, from the perspective of the educational function, is that the composition of AR texts positions the user as the primary knower role, requiring the user to become the source of the AR text. The description of the Creator in Table 1 can explain why the multimodal composition is more challenging to students than the multimodal consumption, and why in the curriculum design, the multimodal viewing literacy should be developed prior to the multimodal composition/representing literacy.

We suggest teachers to use the conceptualisation of the educational function as the guide during their assessment of AR apps for selecting the potential teaching material. To address the question commonly ignored—How does the selected AR app assist the achievement of the learning goal?—teachers should ask themselves the following three sub-questions: What user roles can students play in the interactivity with the selected AR app? By playing these user roles, can students demonstrate the targeting educational function/s in AR-mediated learning? By demonstrating the targeting educational function/s, can students develop the targeting multimodal literacies? These three sub-questions unpack the logic in the meta-reflection about the effectiveness of the selected AR app for AR-mediated learning. Furthermore, these three sub-questions shape the path for teachers to link the selection of AR apps to the achievement of the targeting learning goal in AR-mediated learning.

6. Conclusions

This study has conceptualised the user–app interactivity for the AR-mediated learning. Our conceptualisation includes the identification of the four user roles, the discussion of the resonance between each user role and the prior multimodal experience, and the description of four correspondent educational functions. The conceptualisation provides a consistent, systematic framework for the evaluation of AR apps and AR technology from the Social Semiotic perspective, which views AR as the semiotic discourse. Meanwhile, conceptualisation positions the description and evaluation of AR technology in the educational context, equipping teachers with accessible lenses to make mindful designs and deliveries of AR-mediated learning.

One key limitation of our study was the range of AR apps evaluated and examined. Due to the budget constraint, we could only evaluate and examine AR apps that were free and relatively low-cost. As this study was part of a larger scale project on the use of AR for primary language and literacy education, the apps we chose were also more tailored towards this disciplinary focus that could fit into the participating schools' curricular agenda. Hence, the AR apps featured in this study were not exhaustive and we are not able to guarantee its availability by the time the paper is published or read. Like in all other emerging technologies, the AR apps may be modified by the respective developers in the near future after the publication of this study. Another limitation is the small sample size of six school teachers. Our qualitative observation and analysis of these six teachers' interactive choices made in their user–app interactivity will pave the way for future research to accommodate new AR apps for exploration. To enhance the generalizability of the observation and analysis results, future research is suggested to recruit a larger size of participants, which can potentially instantiate new forms of user–app interactivity.

Future studies are suggested to co-design lesson sequences and to integrate the ongoing exploration of AR apps to create classroom impact. Such pedagogical intervention has proven to be more pragmatic and contextualized to transform teaching and learning [71]. Nevertheless, for teachers who have their first foray into AR-mediated learning, a study like ours is the first step in increasing their awareness of AR-mediated learning before further interventions can be planned prescriptively.

Author Contributions: Conceptualization, T.W. and L.T.; methodology, T.W. and L.T.; validation, T.W. and L.T.; formal analysis, T.W. and L.T.; investigation, T.W. and L.T.; resources, L.T.; data curation, L.T.; writing—original draft preparation, T.W. and L.T.; writing—review and editing, T.W. and L.T.; visualization, T.W. and L.T.; supervision, L.T.; project administration, L.T.; funding acquisition, L.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Primary English Teaching Association Australia (PETAA), grant number 20141.61295. The APC was funded by Japan Society for the Promotion of Science (JSPS), grant number 22K13178.

Institutional Review Board Statement: This research was conducted as part of a larger scale project "Language and Literacy Learning and Teaching with Augmented Reality" funded by the Primary English Teaching Association Australia (PETAA) (2019–2022). The study was approved by the Human Research Ethics Committee (H13489) and complied with the guidelines by NSW State Education Research Applications Process (SERAP number 2019482).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Not applicable.

Acknowledgments: We would like to thank the Japan Society for the Promotion of Science (JSPS) for funding the APC of this publication, grant number 22K13178.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Description of AR apps.

Viewer-Based AR Apps	Brief Description
AR Makr (https://www.armakr.app/ accessed on 4 February 2020) version 1.12; iOS 12.4 or later	AR Makr is a toolbox for sketching and scanning the user's own creations and transferring the creation from 2D to 3D virtual objects.
AR Moon (https://apps.apple.com/us/app/ar-moon-explore-solar-system/id1287945174 accessed on 4 February 2020) version 1.1.4; iOS 11.0 or later	AR Moon is an AR app for the exploration of the surface of the Moon.
ABC AR Space Discovery (https://create.withgoogle.com/inspiration/abc-ar-space-discovery-app accessed on 4 February 2020) version 1.4.2	ABC AR Space Discovery is an AR app for the exploration of the Space as an astronaut or a pilot.
BBC Civilisation AR (https://www.bbc.co.uk/taster/pilots/civilisations-ar accessed on 4 February 2020) version 2.1; iOS 11.0 or later.	BBC Civilisation AR brings art and culture from across the world to the user, exploring the history behind masterpieces.
FarmAR (https://apps.apple.com/us/app/farmar-app/id1407863152 accessed on 4 February 2020) version 3.1.2; iOS 11.1 or later	FarmAR app allows the users to create their own virtual farms and introduces information about soil conditions, crops and animals to the users.
Figment AR (https://viromedia.com/figment accessed on 4 February 2020) version 1.0; Android 7.0 & up version 1.3 (1); iOS 11.0 or later)	Figment AR is an AR app allowing the user to create imaginative scenes and to capture them to share with friends.
Ko'Ko's Curse (https://futureofstorytelling.org/project/ko-ko-s-curse accessed on 4 February 2020) version 1.0; Android 7.0 & up	Ko'Ko's Curse is an interactive fantasy tale app for the teaching about the environment information to children users.
Metaverse (https://studio.gometa.io/landing accessed on 5 February 2020) version 4.0.14; Android 4.4 & up	Metaverse is a creative AR app allowing the user to create his or her own AR experience.

Table A1. *Cont.*

Viewer-Based AR Apps	Brief Description
Storyfab (https://apps.apple.com/au/app/storyfab-movie-studio/id1112571886 accessed on 5 February 2020) version 1.22; iOS 9.0 or later	Storyfab is an AR app allowing the user to create his or her own short films.
Thyng (http://thyng.com/ accessed on 5 February 2020) version 5.0; Android 7.0 & up version 5.0.1; iOS 11.3 or later	Thyng is an AR app allowing the user to create his or her own virtual objects and to transfer the virtual objects into any physical environment.
Quiver (Education Starter Pack–Earth) (https://www.quivervision.com/ accessed on 5 February 2020) version 5.4; Android 4.1 & up	Quiver (Earth) brings the Earth to the user and allows the user to transfer the 3D Earth into any physical environment.
Quiver Mask (https://www.quivervision.com/products/apps/quiver-masks accessed on 5 February 2020) version 1.4.2; iOS 9.0 or later	Quiver Masks is an AR app that combines the AR coloring pages with face-tracking to allow the user to design and decorate his or her own masks and hats.
WDR AR 1933–1945 (https://apps.apple.com/us/app/wdr-ar-1933-1945/id1446878251 accessed on 5 February 2020) version 1.0; iOS 12.0 or later	WDR AR 1933-1945 is an AR app that introduces experience of children and teenagers in the Third Reich to users.
Wonderscope (Clio) (https://preloaded.com/work/clios-cosmic-quest/ accessed on 5 February 2020) version 1.18; iOS 11.0 or later	Wonderscope (Clio) is a voice-driven, interactive AR app for users to explore space.

References

1. United Nations. Sustainable Development Goals Report. 2022. Available online: https://unstats.un.org/sdgs/report/2022/The-Sustainable-Development-Goals-Report-2022.pdf (accessed on 29 May 2020).
2. Azuma, R. A survey of augmented reality. *Presence Teleoperators Virtual Environ.* **1997**, *6*, 355–385. [CrossRef]
3. Dede, C.; Jacobson, J.; Richards, J. Introduction: Virtual reality, augmented, and mixed realities in education. In *Virtual, Augmented, and Mixed Realities in Education*; Liu, D., Dede, C., Huang, R., Richards, J., Eds.; Springer: Berlin/Heidelberg, Germany, 2017; pp. 1–16. [CrossRef]
4. Sat, M.; Ilhan, F.; Yukselturk, E. Comparison and evaluation of augmented reality technologies for designing interactive materials. *Educ. Inf. Technol.* **2023**. [CrossRef]
5. Wang, M.; Callaghan, V.; Bernhardt, J.; White, K.; Peña-Rios, A. Augmented reality in education and training: Pedagogical approaches and illustrative case studies. *J. Ambient Intell. Humaniz. Comput.* **2018**, *9*, 1391–1402. [CrossRef]
6. Diegmann, P.; Schmidt-Kraepelin, M.; Eynden, S.; Basten, D. Benefits of Augmented Reality in educational environments—A systematic literature review. *Wirtsch. Proc.* **2015**, *2015*, 103.
7. Milgram, P.; Takemura, H.; Utsumi, A.; Kishino, F. Augmented Reality—A class of displays on the reality-virtuality continuum. In *Telemanipulator and Telepresence Technologies*; SPIE: Bellingham, WA, USA, 1995; Volume 2351, pp. 282–292.
8. Belda-Medina, J.; Calvo-Ferrer, J.R. Integrating augmented reality in language learning: Pre-service teachers' digital competence and attitudes through the TPACK framework. *Educ. Inf. Technol.* **2022**, *27*, 12123–12146. [CrossRef]
9. Agency Mabu. Understanding the Types of Augmented Reality. 2019. Available online: https://www.agencymabu.com/types-augmented-reality/ (accessed on 24 March 2020).
10. Garzón, J.; Pavón, J.; Baldiris, S. Systematic review and meta-analysis of augmented reality in educational settings. *Virtual Real. J. Virtual Real. Soc.* **2019**, *23*, 447–459. [CrossRef]
11. Parmaxi, A.; Demetriou, A.A. Augmented reality in language learning: A state-of-the-art review of 2014–2019. *J. Comput. Assist. Learn.* **2020**, *36*, 861–875. [CrossRef]
12. Pellas, N.; Fotaris, P.; Kazanidis, I.; Wells, D. Augmenting the learning experience in primary and secondary school education: A systematic review of recent trends in augmented reality game-based learning. *Virtual Real. J. Virtual Real. Soc.* **2019**, *23*, 329–346. [CrossRef]
13. Zhang, J.; Li, G.; Huang, Q.; Feng, Q.; Luo, H. Augmented Reality in K–12 Education: A systematic review and meta-Analysis of the literature from 2000 to 2020. *Sustainability* **2022**, *14*, 9725. [CrossRef]

14. Bonner, E.; Reinders, H. Augmented and virtual reality in the language classroom: Practical ideas. *Teach. Engl. Technol.* **2018**, *18*, 33–53.
15. Palaigeorgiou, G.; Griva, E.; Raftogianni, P.D.; Toronidou, M. Improving vocabulary acquisition in a second/foreign language with a mixed reality environment and a drone. In *European Conference on E-Learning*; Ntalianis, K., Andreatos, A., Sgouropoulou, C., Eds.; Academic Conferences Ltd.: London, UK, 2018; pp. 447–455.
16. Tsai, C.C. A comparison of EFL elementary school learners' vocabulary efficiency by using flashcards and augmented reality in Taiwan. *New Educ. Rev.* **2018**, *51*, 53–65. [CrossRef]
17. Dominguez, E. Teaching media literacy: Harnessing the charm of storytelling with NUSHU. *Child. Educ.* **2019**, *95*, 44–48. [CrossRef]
18. Green, M.; McNair, L.; Pierce, C.; Harvey, C. An investigation of augmented reality picture books: Meaningful experiences or missed opportunities? *J. Educ. Multimed. Hypermedia* **2019**, *28*, 357–380.
19. Redondo, B.; Cózar-Gutiérrez, R.; González-Calero, J.A.; Sánchez Ruiz, R. Integration of augmented reality in the teaching of English as a Foreign Language in Early Childhood Education. *Early Child. Educ. J.* **2020**, *48*, 147–155. [CrossRef]
20. Vedadi, S.; Abdullah, Z.B.; Cheok, A.D. The effects of multisensory augmented reality on students' motivation in English language learning. In Proceedings of the 2019 IEEE Global Engineering Education Conference (EDUCON), Dubai, United Arab Emirates, 8–11 April 2019; pp. 1079–1086.
21. Safar, A.H.; Al-Jafar, A.A.; Al-Yousefi, Z.H. The effectiveness of using augmented reality apps in teaching the English alphabet to kindergarten children: A case study in the State of Kuwait. *EURASIA J. Math. Sci. Technol. Educ.* **2016**, *13*, 417–440. [CrossRef]
22. Raeburn, G.; Welton, M.; Tokarchuk, L. Developing a play-anywhere handheld AR storytelling app using remote data collection. *Front. Comput. Sci.* **2022**, *4*, 927177. [CrossRef]
23. Tan, L.; Chik, A. *Between Worlds: Developing Students' Multimodal Literacies Using Augmented Reality*; Primary English Teaching Association Australia: Newtown, Australia, 2022.
24. Tan, L.; Wang, T. Developing spatial literacy using augmented reality. In *Using Innovative Literacies to Develop Leadership and Agency: Inspiring Transformation and Hope*; Pinhasi-Vittorio, L., Ben-Yose, E., Eds.; IGI Global: Hershey, PA, USA, 2023; pp. 100–123.
25. Bower, M. Technology-mediated learning theory. *Br. J. Educ. Technol.* **2019**, *50*, 1035–1048. [CrossRef]
26. Mills, K.A.; Unsworth, L. Multimodal Literacy. In *Oxford Research Encyclopedia of Education*; Noblit, G., Ed.; Oxford University Press: New York, NY, USA, 2017. [CrossRef]
27. Southgate, E.; Smith, S.P.; Cheers, H. *Immersed in the Future: A Roadmap of Existing and Emerging Technology for Career Exploration*; Paper Series Number 3; DICE Research: Newcastle, UK, 2016.
28. Kress, G. *Literacy in the New Media Age*; Routledge: London, UK, 2003.
29. Mills, K.A.; Unsworth, L.; Exley, B. Sensory literacies, the body, and digital media. In *Handbook of Writing, Literacies, and Education in Digital Cultures*; Mills, K.A., Stornaiuolo, A., Smith, A., Pandya, J.Z., Eds.; Routledge: London, UK, 2018; pp. 26–36.
30. Friend, L.; Mills, K.A. Towards a typology of touch in multisensory makerspaces. *Learn. Media Technol.* **2021**, *46*, 465–482. [CrossRef]
31. LaViola, J.; Kruijff, E.; McMahan, R.P.; Bowman, D.A.; Poupyrev, I. *3D User Interfaces: Theory and Practice*, 2nd ed.; Addison-Wesley: Boston, MA, USA, 2017.
32. Mills, K.A.; Unsworth, L.; Scholes, L. *Literacy for Digital Futures: Mind, Body, Text*; Routledge: London, UK, 2023.
33. Toh, W.; Lim, F.V. Using video games for learning: Developing a metalanguage for digital play. *Games Cult.* **2021**, *16*, 583–610. [CrossRef]
34. Mills, K.A. *Literacy Theories for the Digital Age: Social, Critical, Multimodal, Spatial, Material and Sensory Lenses*; Multilingual Matters: Bristol, UK, 2016.
35. O'Halloran, K. Systemic functional-multimodal discourse analysis (SF-MDA): Constructing ideational meaning using language and visual imagery. *Vis. Commun.* **2008**, *7*, 443–475. [CrossRef]
36. Van Leeuwen, T. *Introducing Social Semiotics*; Routledge: London, UK, 2005.
37. Painter, C.; Martin, J.R.; Unsworth, L. *Reading Visual Narratives: Image Analysis of Children's Picture Books*; Equinox: Sheffield, UK, 2013.
38. Zhao, S.; van Leeuwen, T. Understanding semiotic technology in university classrooms: A social semiotic approach to PowerPoint-assisted cultural studies lectures. *Classr. Discourse* **2014**, *5*, 71–90. [CrossRef]
39. Djonov, E.; van Leeuwen, T. Social media as semiotic technology and social practice: The case of ResearchGate's design and its potential to transform social practice. *Soc. Semiot.* **2018**, *28*, 641–664. [CrossRef]
40. Flowerdew, J.; Richardson, J.E. *The Routledge Handbook of Critical Discourse Studies*, 1st ed.; Routledge: London, UK, 2018. [CrossRef]
41. Djonov, E.; Tseng, C.-I.; Lim, F.V. Children's experiences with a transmedia narrative: Insights for promoting critical multimodal literacy in the digital age. *Discourse Context Media* **2021**, *43*, 100493. [CrossRef]
42. Halliday, M.A.K. *An Introduction to Functional Grammar*; Edward Arnold: London, UK, 1985.
43. Halliday, M.A.K. *An Introduction to Functional Grammar*, 2nd ed.; Edward Arnold: London, UK, 1994.
44. Halliday, M.A.K.; Matthiessen, C.M.I.M. *An Introduction to Functional Grammar*, 3rd ed.; Routledge: London, UK, 2004.
45. Martin, J.R. *English Text System and Structure*; Benjamins: Amsterdam, The Netherlands, 1992.
46. Halliday, M.A.K. *Language as Social Semiotic: The Social Interpretation of Language and Meaning*; Edward Arnold: London, UK, 1978.

47. Kress, G.; van Leeuwen, T. *Reading Images: The Grammar of Visual Design*, 2nd ed.; Routledge: London, UK, 2006. [CrossRef]
48. Kress, G.; van Leeuwen, T. *Reading Images: The Grammar of Visual Design*, 3rd ed.; Routledge: London, UK, 2021.
49. Wang, T. Communicating Climate Change: A Social Semiotic Perspective on Activist Documentaries. Ph.D. Thesis, The University of Sydney, Sydney, Australia, 2020.
50. Wang, T. Developing environmental literacy through eco-documentary-mediated climate change education: A metalanguage toolkit with multimodal approaches. In *Enhancing Education through Multidisciplinary Film Teaching Methodologies*; DeHart, J., Ed.; IGI Global: Hershey, PA, USA, 2023; pp. 122–141.
51. Haas, C. *Writing Technology: Studies on the Materiality of Literacy*; Laurence Erlbaum Associates: Mahwah, NJ, USA, 1996.
52. Bezemer, J.; Kress, G. Touch: A Resource for Making Meaning. *Aust. J. Lang. Lit.* **2014**, *37*, 77–85. [CrossRef]
53. Zhao, S.; Unsworth, L. Touch design and narrative interpretations. In *Apps, Technology and Younger Learners: International Evidence for Teaching*; Kucirkova, N., Falloon, G., Eds.; Routledge: London, UK, 2017; pp. 89–102.
54. Jewitt, C.; Price, S.; Mackley, K.L.; Yiannoutsou, N.; Atkinson, D. *Interdisciplinary Insights for Digital Touch Communication*; Springer Nature: London, UK, 2020.
55. Jewitt, C.; Mackley, K.L.; Price, S. Digital touch for remote personal communication: An emergent sociotechnical imaginary. *New Media Soc.* **2021**, *23*, 99–120. [CrossRef]
56. Bolter, J.; Engberg, M.; MacIntyre, B. Media studies, mobile Augmented Reality, and interaction design. *Interactions* **2013**, *20*, 36–45. [CrossRef]
57. Green, B.; Beavis, C. Literacy education in the age of new media. In *International Handbook of Research on Children's Literacy, Learning, and Culture*; Hall, K., Cremin, T., Comber, B., Moll, L.C., Eds.; John Wiley & Sons Ltd.: Hoboken, NJ, USA, 2013; pp. 42–53.
58. Health, S.B. The handoff play in literacy learning. In *International Handbook of Research on Children's Literacy, Learning, and Culture*; Hall, K., Cremin, T., Comber, B., Moll, L.C., Eds.; John Wiley & Sons Ltd.: Hoboken, NJ, USA, 2013; pp. 184–198.
59. Jang, S.; Vitale, J.M.; Jyung, R.W.; Black, J.B. Direct manipulation is better than passive viewing for learning anatomy in a three-dimensional virtual reality environment. *Comput. Educ.* **2017**, *106*, 150–165. [CrossRef]
60. Lankshear, C.; Knobel, M. *New Literacies: Everyday Practices and Classroom Learning*, 2nd ed.; Open University Press: Maidenhead, UK, 2006.
61. Davies, J.; Merchant, G. *Web 2.0 for Schools: Learning and Social Participation*; Peter Lang: New York, NY, USA, 2009.
62. Milgram, P.; Kishino, F. A taxonomy of mixed reality visual displays. *IEICE Trans. Inf. Syst.* **1994**, *77*, 1321–1329.
63. Blevins, B. Teaching digital literacy composing concepts: Focusing on the layers of Augmented Reality in an era of changing technology. *Comput. Compos.* **2018**, *50*, 21–38. [CrossRef]
64. Fernandez, M. Augmented virtual reality: How to improve education systems. *High. Learn. Res. Commun.* **2017**, *7*, 1–15. [CrossRef]
65. Meletiou-Mavrotheris, M.; Carrilho, A.R.; Charalambous, C.; Mavrou, K.; Christou, C. Teacher training for 'augmented reading': The living book approach and initial results. *Educ. Sci.* **2020**, *10*, 144. [CrossRef]
66. Aso, B.; Navarro-Neri, I.; García-Ceballos, S.; Rivero, P. Quality requirements for implementing Augmented Reality in heritage spaces: Teachers' perspective. *Educ. Sci.* **2021**, *11*, 405. [CrossRef]
67. Liang, W.J.; Lim, F.V. A pedagogical framework for digital multimodal composing in English language classroom. *Innov. Lang. Learn. Teach.* **2021**, *15*, 306–320. [CrossRef]
68. Tan, L.; Thomas, A.; Chik, A.; Novianti, N.; Jones, G. Teaching language, literacy and literature using augmented reality: An introduction to AR apps and their potential for classroom use. *PETAA Pap.* **2020**, *220*, 1–11.
69. Berry, M. Systemic Linguistics and Discourse Analysis: A multi-layered approach to exchange structure. In *Coulthard Montgomery*; Routledge: London, UK, 1981; pp. 120–145.
70. Berry, M. Towards layers of exchange structure for directive exhanges. *Network* **1981**, *2*, 23–32.
71. Reinking, D.; Bradley, B. *Formative and Design Experiments: Approaches to Language and Literacy Research*; Teachers College Press: New York, NY, USA, 2008.