The JALT CALL Journal
ISSN 1832-4215, 21(3), 102668 (2025)
https://doi.org/10.29140/jaltcall.v21n3.102668

Castledown

Comparing augmented reality and quick response code versions of a location-based task for foreign language learning

Samuel Taylor
Kyushu Sangyo University
taylor@mail.kyusan-u.ac.jp

Location-based Augmented Reality (AR) technology has the potential to transform language education by enabling situated, authentic, and collaborative learning. Yet, while some research has found that AR facilitates learning, other research has found that AR increases cognitive load, is distracting, and provides too much information during learning activities. This study compared the student experiences of AR and Quick Response (QR) code versions of a campus-based English language activity at a Japanese university. Exploratory Factor Analysis of 483 students' post-task self-report measures of motivation, value, and difficulty identified a reliable scale for motivation. Results showed that AR made the task more motivating and encouraged collaboration. However, AR made smartphone technology more difficult to use and perceived performance to be lower. These challenges were attributed to design issues regarding learner interaction with learning content, such as redundant and transient information. The findings suggest that AR is best suited to tasks that prioritize collaboration and the co-construction of knowledge. In designing AR tasks, and Mobile Assisted Langauge Learning materials in general, reducing cognitive load and improving performance are key issues that can be addressed with carefully designed materials that have clear structure, concise content, and guidance on how to interact with them effectively.

Keywords: Augmented reality, mobile assisted language learning, cognitive load, collaborative learning

Introduction

Augmented Reality (AR) is a tool of educational technology that supports (Boettcher, 2007) and transforms (Hoopingarner, 2009) the learning experience by enhancing the efficiency, quality, and character of learning (An & Oliver, 2020). In the current study AR is defined as the leveraging of smartphone capabilities to provide an additional layer of information that is interactive in real time (Dunleavy & Dede, 2013; Salmon & Nyhan, 2013). Research on AR use in educational settings has identified benefits related to motivation, content understanding, long-term memory retention, physical task performance, and collaboration (Radu, 2014). In regard to foreign language learning, AR has positive benefits when combined with pedagogically suitable learning tasks (Hockly, 2019; Taylor, 2018), and can improve motivation and engagement, learning performance, interaction and collaboration, and opportunities for authentic language use (Parmaxi & Demetriou, 2020). Conversely, AR has also been found to be detrimental to learning by requiring high attentional demands, and through technological issues, usability difficulties, ineffective classroom integration, and learner differences (Buchner & Kerres, 2021; Radu, 2014). In particular, contradictory findings exist regarding cognitive load. While some research has found that AR reduces cognitive load, facilitating learning, other research has found that AR increases it, is distracting and provides too much information during learning activities (Buchner & Kerres, 2021). Evidence from comparative studies of AR and other media is also varied. While most studies report lower or equal cognitive load with higher performance when compared to other conditions (see Lai et al., 2018), there is evidence of cognitive load and performance being high, and of AR showing worse performance due to cognitive load (Buchner & Kerres, 2021; Skulmowski & Rey, 2020).

AR-enabled location-based language learning activities involve overlaying content on the physical environment and allow learners to move their use of language to a meaningful context (Holden & Sykes, 2011). Such activities facilitate the creation of a learning community (Dunleavy et al., 2009; Lee & Park, 2019), encourage exploration and authentic learning (Pombo et al., 2017), and favor collaborative and negotiated problem-solving in useful activities (Morrison et al., 2009; Perry, 2021) valued by learners (Liu et al., 2010). Quick Response (QR) code-enabled location-based activities (see Law & So, 2010; Rikala & Kankaanranta, 2012) involve content being accessed through the physical environment but not overlaid on to it. These activities have also been found to enhance personalized, situated, authentic, and collaborative learning (Rikala & Kankaanranta, 2012).

Expanding on the promise of location-based learning to enable authentic, situated, and collaborative learning, a campus-based learning activity that requires groups of students to understand directions to, and information about, campus locations facilitates authentic learning in three ways. First, it allows students the chance to interact naturally to achieve a particular communicative goal. Second, it relates to real world target tasks. Third, it engages students in a way that they are interested in its topic and its purpose, and understand its relevance (Guariento & Morley, 2001). Furthermore, location-based learning embodies principles of situated learning, which can be defined as 'learning knowledge and skills in contexts that reflect the way the knowledge will be useful in real life' (Collins, 1988: 2), through the placing of groups of students in the authentic context of the university campus. In this authentic context, the fundamental premise upon which situated learning theory rests (Herrington & Oliver, 2000), opportunities for the construction of knowledge through interactions with others creates conditions for learning (Lave & Wenger, 1991). Collaboration is facilitated by AR-enabled learning through what Morrison et al. (2009) describe as 'place-making', by requiring constant referencing to the physical environment and allowing bodily configurations that establish common ground for group interaction.

The validity and suitability of AR-enabled location-based activities to language teaching and learning can be considered in relation to the degree to which motivating, valuable, and suitably difficult activities (Hwang et al., 2013, McQuiggan et al, 2015) engage learners in the cognitive, affective, social, and behavioral aspects of meaningful learning (Conceicao & Howles, 2021; Lee & Park, 2019).

Motivation results in high-quality learning (Ryan & Deci, 2000) and is influenced by emotions, such as interest and frustration (Conceicao & Howles, 2021). Morrison et al. (2009) used items from the Intrinsic Motivation Inventory (IMI) to measure differences in student motivation when playing an AR-enabled paper map version and a non-AR 2D digital map version of a location-based game. Student interest was higher in the non-AR version than the AR version, suggesting that affective learning was better served by not using AR. Behavioral learning was also differentiated, with users of the AR version being more concentrated on the technology and more focused on the game. Non-AR version users rated their task performance higher than the AR version users, which suggests that AR did not improve cognitive learning. Morrison et al.'s (2009) findings are in contrast to Lai et al. (2018), who found AR to be more motivating and to require less mental effort than conventional multimedia.

Value is linked to motivation in that motivated learners are more likely to perceive a mobile learning environment as suitable for a language learning task. In this way, value can be seen as reflecting learners' intention to use mobile devices in language learning (Sun & Gao, 2020). A similar relationship between motivation and value was found by Liu et al. (2010), when using a survey to investigate learners' perceived value of a 2D barcode, handheld, AR-supported English learning environment. Learners found the environment useful and motivating for assisting learning. Eight survey items targeted sources of value and showed that for learners, value related to the affective dimensions of interest and motivation, the cognitive dimension of specific skills development (listening, speaking, reading, and writing), and the behavioral dimension of the application of knowledge to real life.

Difficulty experienced by learners in AR-enabled learning environments has been investigated in a variety of ways. In regard to a location-based AR mobile game, Pombo et al. (2017) measured student perceptions of task difficulty using ratings of easiness of task completion, as well as participant observation and focus groups. However, most research focuses on the cognitive dimension of load. AR is said to reduce cognitive load through object annotation (Santos et al., 2014). This is explained using principles from Multimedia Theory regarding the use of virtual words to annotate physical objects (Mayer & Merino, 2003).

AR increases cognitive load if part of a complex series of tasks. Dunleavy et al. (2009) found the synthesis of skills involved in a location-based educational activity that required students to perform geo-spatial navigation, collaborative problem solving, handheld manipulation, and mathematics and literacy content comprehension, to be the cause of cognitive overload.

Other sources of cognitive overload may concern what Sweller (2011) terms the redundancy effect and the transient information effect. Marker-based AR may involve the addition of redundant information that interferes with learning. For example, watching a video overlaid onto an AR marker, a student may be distracted by having to position their device at an appropriate distance from and angle to the AR marker. The redundant information processed when viewing content increases mental effort, or extraneous cognitive load, which decreases the capacity for the processing of the mental load required for learning content comprehension (Sweller, 2011). Furthermore, AR may make information transient as it can only be accessed at certain locations or by viewing certain physical objects. This loss of permanency and ease of access is said to increase mental effort and subsequently cognitive load, especially with long segments of information (Wong et al., 2012). In a long segment of information, whatever is presented at the end will be hard for working memory to process as working memory have already been exhausted by the previously presented information.

Aims and hypotheses

This study investigates differences between student experiences of AR and QR code versions of a location-based task and what these differences may reveal about the validity and suitability of AR as a tool for foreign language instruction. Comparison of the two modes is potentially useful to instructors if it reveals sources of motivation, value, and difficulty, allowing for tasks to be redesigned to improve student performance and learning (Dunleavy & Dede, 2013). This approach also adds to the limited body of research on comparisons of different modes of delivering Mobile Assisted Language Learning (MALL) to provide insight for curriculum and instructional designers on mobile interface design (Çakmak, 2019) to enable meaningful learning (Chwo et al., 2018).

It is hypothesized that the location-based task will be highly motivating for students due to its authentic context, and that it will also be perceived as highly valuable, or useful for learning. Motivation and value are expected to correlate highly. Difficulty is expected to not correlate to motivation or value. In comparing AR and QR code versions, it is hypothesized that both will facilitate authentic, situated, and collaborative learning. Also, that the AR version will be more motivating for students due to the impact of the real time augmentation of the physical environment with virtual content. It is also expected that perceptions of the value of the AR version of the task will be higher than those of the QR code version. In regard to difficulty, it is anticipated that, despite being more motivating, the AR version of the task will be perceived as more mentally taxing due to both the redundancy effect and the transient information effect.

Materials and methods

The digital campus-based task was conducted, and a post-task survey administered to a total of 522 first to third year students in 25 classes by twelve instructors of compulsory English listening and speaking classes at a large private university in southwest Japan. Participants' English proficiency ranged from CEFR A2 to CEFR B1 (Council of Europe, 2001). The student body was technology-rich: nearly all students had a smartphone, and were generally very adept at using it, and the university campus had freely accessible wi-fi. Demographic data, such as gender and age, was not recorded.

The campus-based activity presented information about campus locations. It was designed to be used in a 90-minute class, with expected learning outcomes being the comprehension of basic location and direction vocabulary, and descriptions of campus facilities. Learning content was written and video recorded by students, who were not included in the current study, and included practical information, personal opinions or preferences, comprehension check questions, and directions to another campus location. A sample of the learning content is provided in Table 1.

Table 1. Example learning content from the digital campus-based task

Campus location	Description	Worksheet question/answer	Directions to next location
Library	I want to introduce my favorite place on campus. This is the library. It is large and is built with red bricks. The library has about 740,000 books and many technical books in various fields. And there are many audio-visual materials and electronic publications, such as, CD-ROMs. Enter, then pass the room where you can eat and drink on the first floor. Walk up the nearest stairs. Then, you can see the self-study room on the second floor. This is my favorite place.	How many books does the library have? / 740,000	(Arteria canteen) The next treasure is between building 15 and the library. Find the entrance near some vending machines, close to the Kakiemon Youshiki Gama.
Arteria canteen	This is Arteria, one of the canteens on campus. It has a lot of chairs and tables so you can come here to eat meals, but also just to chat with friends in time between classes. I use this building almost every day. I am a life science major, so I use this building for many of my classes. Inside the dining hall, there is an ice cream vending machine which is probably the only one on campus. On a hot day, if you want to eat ice cream, you can choose from seventeen flavors. This will surely make you happy and smile. I think this will be your favorite place on campus.	How many flavors of ice cream are there in the ice cream vending machine? / Seventeen	(Art museum) The next treasure is located between the North gate and the Central Hall. There are many works of art on display here. Look for the red windmill and colorful poles.

The learning content was recorded in one video and was either uploaded to the HP Reveal app and overlaid on AR markers, which were images associated with a pirate-themed treasure hunt or uploaded to YouTube and linked to QR codes using the video URLs. Both versions of the task used the same content and the same images on the physical markers that were placed at locations around campus. While content was accessed in the AR version by aiming a smartphone camera at the images on the markers, content in the QR code version was accessed by aiming at the QR codes placed underneath each image on the physical markers. Each place with a marker was incorporated into a loop by its video containing directions to find the next marker. To complete the activity, students had to find all the markers, or treasure, in the loop, watch the videos, and answer the questions provided on a worksheet. Classes of students did either the AR or QR version of the activity. A total of 294 students in 16 classes did the AR version of the activity and a total of 228 students in nine classes did the QR code version of the activity. Students were put into pairs or groups of three and given a worksheet with activity instructions, questions about each location, and a map of the campus. Once students completed the activity by visiting all campus locations and answering all the questions, they returned to the classroom and completed the anonymous online survey.

The constructs of motivation, value and difficulty were measured as follows. The interest and value scales from the IMI (Ryan, 1982) were used to measure motivation and value as the IMI is usually adapted for specific contexts and used in combination with other measures (see Morrison et al., 2009; Sun & Gao, 2020). The NASA Task Load Index (TLX) (Hart & Staveland, 1988) was used to measure difficulty. It is a measure of workload that includes the dimension of cognitive load and has been used in the evaluation of both AR and mobile technologies (Hart, 2006). Five items were taken from each scale. The researcher, in collaboration with a colleague, then added a further fifteen items that aimed to measure possible sources of motivation, value, and load specific to the conditions of the learning task (Skulmowski & Rey, 2020). These sources were determined to be cognitive (comprehension of learning content), affective (enjoyment of AR), behavioral (use of smartphones, learning information about campus), social (working with classmates), and situated (moving around campus) in nature (Conceicao & Howles, 2021; Lee & Park, 2019). The final survey contained 30 items, and used a 4 point Likert scale to enable dichotomous results to be obtained. Table 2 displays the survey themes, the items taken from existing measures, and the items added to investigate possible sources of motivation, value, and load.

Table 2. Survey item themes, item sources, item codes, and items

Theme	Item source	Item code	Item
Motivation	IMI Interest	M1	I was able to engage in the activity naturally.
		M2	I enjoyed this activity very much.
		M3	I would like to do this activity again.
		M4	I don't think that others will find this activity interesting.
		M5	I found this activity boring.
Affective motivation	Original	M6	Watching videos was fun.
Situated motivation		M7	Searching for locations was fun.
Behavioral motivation		M8	Learning about campus was fun.
Behavioral motivation		M9	Using my smartphone was fun.
Social motivation		M10	Working with classmates was fun.
Value	IMI Value	V1	I think this activity was a waste of time.
		V2	I thought this activity was important because by doing it I could improve my English language skills.
		V3	I think this activity is not useful for learning English.
		V4	This activity is an effective way of learning English.
		V5	I believe this activity has value.
Situated value	Original	V6	Moving around the campus was helpful for learning.
Cognitive value		V7	This activity was helpful for learning prepositions.
Cognitive value		V8	This activity was helpful for learning the passive voice.
Behavioral value		V9	I did not learn anything new about the campus.
Social value		V10	This activity was easy to do in co-operation with classmates.
Load	TLX	L1	It was mentally difficult to do the activity.
		L2	It was not physically difficult to do the activity.
		L3	I was frustrated by this activity.
		L4	I had to rush to finish the activity in time.
		L5	I successfully completed the activity.
Behavioral load	Original	L6	I was distracted by my smartphone.
Behavioral load		L7	Smartphone technology was easy to use.
Cognitive load		L8	The videos were too long.
Cognitive load		L9	The videos were difficult to understand or hear.
Situated load		L10	I was not distracted by working outside.

Survey responses were exported to a spreadsheet for analysis. Incomplete survey responses – any response that did not contain responses to all thirty survey items – were deleted. Exploratory factor analysis (EFA) was used to determine whether any themes that informed the design of the survey formed valid and reliable scales. The initial analysis, however, established whether or not the data were evenly distributed. Skewness and kurtosis values were inspected and any value greater than +1 or −1 was judged to indicate abnormal distribution and inappropriateness for inclusion in analysis. Following this, Bartlett's test of sphericity and the Kaiser–Meyer–Olkin Measure of Sampling Adequacy (MSA) were used to determine whether sufficiently large relationships existed within the data set to perform EFA (Howard, 2016). The factor analytic method was maximum likelihood (ML) estimation and parallel analysis, and visual scree plot (VSP) analysis were used together to inform factor retention decisions. Oblique rotation was used as the factors were considered to correlate (Costello & Osbourne, 2005). Decisions regarding the factor loading cutoff were made using the 0.40–0.30–0.20 rule. This states that satisfactory variables load onto their primary factor above 0.40, load onto other factors below 0.30, and demonstrate a difference of 0.20 between their primary and other factor loadings (Howard, 2016).

A reliability analysis examined the internal consistency of the produced factors to determine whether they could be treated as distinct and internally consistent scales. A correlational analysis of the scales and individual survey items investigated the relationship between variables. Correlation coefficients were calculated using Pearson's *r* value as scatter plots indicated linear relationships between variables. Following Plonsky & Oswald (2014), *r* values close to 0.25 were considered weak, close to 0.40 medium, and close to 0.60 large. The scales and individual items were then subjected to independent samples t-tests to check for any statistical significant differences in the student experience of the AR and QR code versions of the activity. The Welch test was used as it assumes unequal variances. The practical significance of mean differences between variables were considered using Cohen's *d* values. L2 field-specific benchmarks of small (d=0.40), medium (d=0.70), and large (d=1.00) were used (Plonsky & Oswald, 2014).

Results

A total of 514 survey responses were received from the 522 students in 25 separate classes. Once 31 incomplete responses were removed, there remained 483 responses for use in analysis: 265 from the AR version, and 218 from the QR code version. The mean for all the items combined was 2.81, and the median was 2.80, indicating that students had a largely positive experience of the task. Variable distributions were generally negatively skewed and platykurtic. High-scoring items, those in the upper quartile (Q3>3.12, IQR=0.64,) showed that students enjoyed and valued working with their classmates (M10, m=3.52; V10, m=3.48), that using smartphones was not distracting (L6, m=3.40), and that the task was engaging (M1, m=3.24) and enjoyable (M2, m=3.17). Conversely, low-scoring items, those in the lower quartile (Q1<2.48), revealed that students felt rushed (L4, m=2.11) and that it was physically difficult to complete the task (L2, m=2.35), that the video content was too long and difficult to understand (L8, m=2.35; L9, m=2.17), and that students were unsure of the degree to which they successfully completed the activity (L5, m=2.42). Table 3 displays the means, standard deviations, skew and kurtosis values, and MSA values for all survey items.

Table 3. Motivation (M), value (V), and difficulty (D) survey scales, items, and item mean (M), standard deviation (SD), skew, kurtosis, and measure of sampling adequacy (MSA) values. N=483

Item code	Item	M	SD	Skew	Kurtosis	MSA
M10 [ab]	*Working with classmates was fun.	3.52	0.72	−1.60	2.33	0.89
V10 [ab]	*This activity was easy to do with my classmates.	3.48	0.68	−1.22	1.33	0.91
L6 [abd]	*I was not distracted by my smartphone.	3.40	0.82	−1.32	1.15	0.78
M1	I was able to engage in the activity naturally.	3.24	0.79	−0.83	0.19	0.94
M2	I enjoyed this activity very much.	3.17	0.88	−0.80	−0.19	0.95
M5	I did not find this activity boring.	3.10	0.90	−0.76	−0.22	0.96
M7	Searching for locations was fun.	3.06	0.94	−0.70	−0.46	0.96
V1	I don’t think this activity was a waste of time.	3.03	0.97	−0.67	−0.60	0.91
L10 [d]	I was not distracted by working outside.	3.03	0.96	−0.65	−0.57	0.90
L3	I was not frustrated by doing the activity.	3.00	0.98	−0.56	−0.81	0.95
V5	I believe this activity has value.	2.99	0.90	−0.57	−0.46	0.95
M9	Using my smartphone was fun.	2.91	0.94	−0.49	−0.65	0.96
V3	I think this activity is useful for learning English.	2.90	0.96	−0.52	−0.69	0.89
M8	Learning about campus was fun.	2.87	0.94	−0.39	−0.78	0.95
L7	The smartphone technology was easy to use.	2.85	0.97	−0.47	−0.74	0.92
V4	This activity is an effective way of learning English.	2.83	0.91	−0.39	−0.63	0.94
V9	I learnt something new about campus.	2.83	0.90	−0.45	−0.50	0.90
V6	Moving around campus was useful for learning.	2.82	0.95	−0.37	−0.80	0.95
M3 [b]	*I want to do this activity again.	2.69	1.09	−0.26	−1.23	0.96
V2	I thought this activity was important because by doing it I could improve my English language skills.	2.66	0.91	−0.23	−0.72	0.95
V7	This activity was useful for learning prepositions.	2.58	0.92	−0.08	−0.80	0.83
V8	This activity was useful for learning the passive voice.	2.58	0.92	−0.07	−0.84	0.84
M6	Watching videos was fun.	2.55	0.96	−0.01	−0.94	0.95
L1 [b]	*It was not mentally difficult to do the activity.	2.50	1.05	−0.01	−1.18	0.90
L5 [bd]	*I successfully completed the activity.	2.42	1.03	0.08	−1.13	0.80
L2 [b]	*It was not physically difficult to do the activity.	2.35	1.05	0.14	−1.18	0.92
L8 [bcd]	*The videos were not too long.	2.35	1.02	0.14	−1.10	0.59

Item code	Item	M	SD	Skew	Kurtosis	MSA
M4 [d]	*I don't think that others will find this activity interesting.	2.22	0.94	0.29	−0.82	0.85
L9 [cd]	*The videos were not difficult to understand or hear.	2.17	0.97	0.41	−0.83	0.62
L4 [cd]	*I didn't have to rush to finish the activity in time.	2.11	1.03	0.51	−0.90	0.59
	Total	2.81	0.48	−0.30	−0.18	0.92

[a] high skewness, [b] high kurtosis, [c] low MSA value, [d] low collinearity, *not included in EFA.

After consideration of the distribution, sample adequacy, sphericity, and collinearity, 19 of 30 items were determined to be factorable. EFA of the 19 retained variables led to six variables being removed for low and cross factor loadings, and a four-factor model being produced. All factor loading variances are displayed in Table 4. Additional fit indices, RMSEA (0.061) and TLI (0.96) values, indicate a good degree of fit for this four-factor model. The first factor, motivation, consisted of four original items and two taken from the IMI Interest scale. The other three factors consisted of original and IMI Value scale items, and distinguished cognitive value from behavioral value.

The reliability analysis indicated that only the first factor formed a reliable scale (Cronbach's α = 0.86). This scale for motivation was used in subsequent analysis. The mean value of the scale was 2.96 (SD=0.69), as shown in Table 5 along with the mean scores of each item in the scale, indicating that the task was satisfying (Ryan & Deci, 2000) and, in particular, because it was situated (m=3.06), authentic (m=2.87), and used mobile technology (m=2.91).

Table 4. Factor loadings and variances of the four-factor result of EFA

Factor	Item theme / source	Item	Factor 1	Factor 2	Factor 3	Factor 4
Motivation	Situated motivation	Searching for locations was fun.	.76			
	Behavioral motivation	Learning about campus was fun.	.73			
	Behavioral motivation	Using my smartphone was fun.	.67			
	IMI interest	I enjoyed doing this activity very much.	.64			
	Affective motivation	Watching videos was fun.	.57			
	IMI interest	I was able to engage in the activity naturally.	.54			
Cognitive value	Original	This activity was useful for learning the passive voice.		.97		
	Original	This activity was useful for learning prepositions.		.80		
Behavioral value	IMI value	I don't think that this activity was a waste of time.			.88	
	IMI value	I think that this activity was useful for learning English.			.66	
	Behavioral value	I didn't learn anything new about campus.			.56	
Value	IMI value	This activity is an effective way of learning English.				.76
	IMI value	I think that this was important to do because I could improve my English abilities.				.65
		Proportion of variance	.23	.15	.13	.11
		Cumulative variance	.23	.38	.51	.62

RMSEA = 0.061, TLI = 0.96.

Table 5. Motivation scale, including scale and individual item means (M) and standard deviations (SD)

Item theme / source	Item code	Item	M	SD
		Motivation scale	2.96	0.69
Situated motivation	M7	Searching for locations was fun.	3.06	0.94
Behavioral motivation	M8	Learning about campus was fun.	2.87	0.94
Behavioral motivation	M9	Using my smartphone was fun.	2.91	0.94
IMI interest	M2	I enjoyed doing this activity very much.	3.17	0.88
Affective motivation	M6	Watching videos was fun.	2.55	0.96
IMI interest	M1	I was able to engage in the activity naturally.	3.24	0.79

Cronbach's α = 0.86.

The correlational analysis revealed statistically significant ($p<0.001$) correlations between the motivation scale and interest (M3, $r=.71$; M5, $r=.61$), and value (V4, $r=.60$; V5, $r=.75$; V6, $r=.68$). There were nine strong correlation coefficients between individual items, the strongest of which suggests that learners valued the activity as a means of learning English (V4 and V5, $r=.77$). The medium and strong correlations are displayed in Table 6.

Table 6. Medium and strong correlation coefficients (Pearson's R) between the motivation scale and individual survey items

	M	M3	M5	M10	V1	V2	V3	V4	V5	V6	V7	V8	V10
Motivation (M)													
IMI interest (M3)	.71												
IMI interest (M5)	.61	.61											
Social motivation (M10)	.59	.41	.43										
IMI value (V1)	.54	.54	.63	.33									
IMI value (V2)	.59	.53	.44	.28	.40								
IMI value (V3)	.32	.34	.48	.22	.64	.36							
IMI value (V4)	.60	.60	.48	.37	.41	.70	.37						
IMI value (V5)	.75	.67	.55	.50	.54	.66	.42	.77					
Situated value (V6)	.68	.54	.38	.40	.38	.57	.25	.60	.67				
Cognitive value (V7)	.49	.41	.26	.25	.23	.50	.21	.49	.49	.49			
Cognitive value (V8)	.52	.40	.27	.27	.24	.54	.18	.52	.52	.46	.84		
Social value (V10)	.51	.36	.36	.62	.32	.31	.22	.38	.45	.39	.28	.29	
TLX (L3)	.50	.52	.56	.28	.53	.34	.39	.37	.43	.36	.21	.24	.25

Note: $p<0.001$ for all correlations.

Independent samples t-tests indicated that the AR version of the activity (M=3.04) was significantly ($p<0.05$) more motivating for students than the QR code version (M=2.87), but with a small effect size ($d=0.25$). Comparing individual survey items, there were significant differences for ten items. The AR version was more enjoyable ($p<0.01$, $d=0.27$), and more fun to do with classmates than the QR code version ($p<0.05$, $d=0.23$), and watching the video content was more fun ($p<0.001$, $d=0.32$) despite the video content for both versions of the task being identical. The AR version was also rated as less boring ($p<0.05$, $d=0.22$), and students wanted to do it again more than they did the QR code version ($p<0.05$, $d=0.23$). The AR version was less mentally demanding ($p<0.05$, $d=0.23$) and less frustrating ($p<0.01$, $d=0.26$) than the QR code version. However, despite favorable ratings for motivation and load, the AR version was perceived as being less successfully completed than the QR code version ($p<0.001$, $d=0.46$). Also, students found the smartphone technology more difficult to use ($p<0.001$, $d=0.56$) and the videos were judged to be too long more than in the QR version (p<0.001, d =0.36). All significant differences between the AR and QR code versions are displayed in Table 7.

Table 7. Survey items with significant difference between the AR and QR code versions, including item themes, means (M), standard deviations (SD), and effect sizes (D)

			AR (n=265)		QR (n=218)		
Theme	Item code	Item	M	SD	M	SD	d
		Motivation Scale	3.04*	0.66	2.87	0.72	0.25
IMI interest	M2	I enjoyed doing this activity very much.	3.27**	0.86	3.04	0.91	0.27
Affective motivation	M6	Watching videos was fun.	2.68***	0.93	2.38	0.96	0.32
IMI interest	M3	I want to do this activity again.	2.80*	1.07	2.55	1.10	0.23
IMI interest	M5	I did not find this activity boring.	3.19*	0.86	2.99	0.94	0.22
Social motivation	M10	Working with classmates was fun.	3.60*	0.67	3.43	0.77	0.23
TLX	L1	It was not mentally difficult to do the activity.	2.61*	1.08	2.37	0.99	0.23
TLX	L3	I was not frustrated by doing the activity.	3.11**	0.94	2.86	1.01	0.26
TLX	L5	I successfully completed the activity.	2.21	0.95	2.67***	1.07	0.46
Behavioral load	L7	The smartphone technology was easy to use.	2.62	1.01	3.13***	0.84	0.56
Cognitive load	L8	The videos were not too long.	2.19	0.98	2.54***	1.03	0.36

Note: *p<0.05, **p<0.01, ***p<0.001.

Discussion

The campus-based activity was an authentic task for the learners in this study. By requiring groups of students to visit and learn about campus facilities, it demonstrated three key features of authentic tasks by having a genuine purpose, using real world subject matter, and engaging students in learning (Guariento & Morley, 2001). Furthermore, the university campus was an authentic, contextual, and engaging learning environment (Perry, 2021), and in this environment, mobile learning was a suitable and authentic means of foreign language instruction. Mobile devices, despite their potential for individualized exploration and distraction, were successfully harnessed to create a motivating, situated, authentic, and collaborative learning environment for classes of students (Chwo et al., 2018; Squire, 2009). Notwithstanding, the authenticity of the task could be extended further by including the learners in content creation. A learner-centered approach to this would involve interactions between teacher and students about the learning targets and content (Lee & Park, 2019) and would also allow for the task to be made part of a larger sequence of tasks integrated into a curriculum.

The motivation to complete the task was shown to derive from contextual factors that were affective, behavioral, and situated in nature, echoing the findings of Lee & Park (2019), who show how location-based AR supports meaningful learning by embodying cognitive, affective, and social dimensions of learning. Motivation correlated strongly with value, which has been shown to have direct associations with students' behavioral intention (Sun & Gao, 2020). The perceived value of a mobile technology-enabled task is important in encouraging high quality cognitive engagement with the learning content (Lee & Park, 2019), and behavioral engagement, such as learner interaction with technology and the application of learning to real world contexts. In particular, this research has shown how the behavioral dimension of learning is an important aspect of motivation in the use of technology (Conceicao & Howles, 2021).

AR was found to be more motivating than conventional methods of content delivery, as found by Lai et al. (2018). Video content being more fun in AR indicates that the magic-like augmentation of the real world with virtual content was enjoyable, and that viewing content overlayed onto the physical environment in real-time, rather than 'in-device', was satisfying. This inherent quality of AR technology may also encourage student-to-student interactions more than QR code use. Although both versions of the task encouraged interaction within groups, and therefore collaborative and negotiated problem-solving (Morrison et al., 2009; Perry, 2021), 'place-setting' (Morrison et al., 2009) may have been established by the AR version of the task by requiring one student to hold a device up to a marker, at an appropriate distance and angle. The writing of answers to questions and note-taking regarding the directions to the next marker then falls to a second student. This dynamic is not produced by the QR code version, in which a student can access a video by momentarily holding up a device to a QR code. Videos can then be watched at a place and time of their own choosing, without requiring the cooperation of another student. However, the QR code version may facilitate greater interaction with the university campus as place-making can lead to a greater focus on the technology and the task at the expense of a lesser focus on the physical environment (Morrison et al., 2009).

By juxtaposing physical locations with virtual information about those locations, AR technology may be improving perception and reducing difficulty by lowering cognitive load in limited working memory (Santos et al., 2014). However, how this differs to the affordances of QR codes is unclear and highlights the limitations of cognitive learning theories for understanding location-based learning. The lower mental demand of AR may be better understood with reference to constructivist learning theory. From this perspective, knowledge is embedded in context, and meaning is constructed through interaction with others (Lave & Wenger, 1991). Via 'place-setting', AR facilitates four of the five conditions proposed by constructivist learning theory to enhance learning: it embeds learning in context, it facilitates social interaction, and it provides multiple perspectives and modes of representation (Dunleavy & Dede, 2013). The task conditions created by AR use may improve learning in this way and consequently lower student perceptions of the difficulty and frustration of doing the task.

Students found exploring the campus taxing and time-consuming. These experiences indicate that the learning content may have been too difficult for students' English proficiency level, causing a high degree of cognitive load (Sweller, 2011). It may also indicate that the cognitive load required for students to interpret the meaning of information about, and directions to campus locations, navigate themselves around campus, and access digital media through AR or QR code markers may have inhibited performance (Dunleavy et al., 2009). Additionally, difficulty may have been affected by the mode of media delivery. Students' interactions with the AR technology led to more negative perceptions of task performance. As the videos were the same length in both versions of the task, this difference indicates cognitive overload. This may be due to a redundancy effect (Sweller, 2011) being caused by the overlaying of virtual content on the physical environment, which required students to hold a smartphone still in a position close enough and at an angle sufficient enough to allow for concentrated viewing without distraction from the physical environment. In designing mobile learning, such cognitive load can be reduced by 'weeding' – a technique described by Mayer & Moreno (2003) as the elimination of interesting but extraneous content. In the context of the current study, this would involve keeping content concise and avoiding including background music or visual filters that may distract from essential processing. Another solution is 'signaling' – the use of signals concerning how to process materials. Signaling can be achieved by stressing key words in speech, coloring key parts of images, and adding headings to text (Mayer & Moreno, 2003). For the activity described in the current study, signaling could be achieved by using the headings 'Facility information', 'Quiz question', and 'Directions to next location' in the learning content.

The difference in perceived video length in the AR and QR-code versions of the task may be explained by a transient information effect (Sweller, 2011) caused by the information about campus locations in the AR-accessed videos being only accessible at the physical AR markers, making it impermanent. This effect has been shown to vary due to the length of the information presented (Wong et al., 2012). It would particularly affect the cognitive load involved in interpreting directions, as these were provided at the end of each video, by which time working memory reserves may have been low. Solutions to this include 'segmenting' – breaking down learning content into learner-controlled segments. The marker-based AR content used in this study could be split into three segments, 'Facility information', 'Quiz question', and 'Directions to next location', each of which could be accessed using labeled markers as and when the learner is ready. 'Pretraining' of signals and segments would further lower cognitive load. Explaining how content is presented and how it may be processed helps learners more effectively process it (Mayer & Moreno, 2003).

Limitations

The current study was conducted at a single Japanese university and involved learning content specific to the university's campus. Therefore, its findings may be contextually bound and not readily generalizable to broader populations. Cautious interpretation of the results regarding their transferability to other cultural or institutional contexts is required. Furthermore, this study's exclusive focus on student perceptions, paired with the absence of learning outcomes data limits the extent to which conclusions can be drawn about pedagogical effectiveness. Future research would benefit from the use of causal evidence of the pedagogical soundness of the task that can be correlated with student perceptions to provide a more complete picture of the tasks' validity and suitability for learning. Measures of learning outcomes, in particular, are better undertaken using data gained from pre- and post-test data (Taylor & Stone, 2023). Such an approach may also compensate for the small effect sizes of the differences found between the student perceptions of the AR and QR code versions of the task.

A further limitation can be found in the survey design. The failure of the 30 items to form three distinct and reliable scales suggests that adaptation of the IMI, and TLX for measuring motivation, value, and difficulty in mobile technology-mediated location-based learning may not be suitable. However, the success of the researcher-created items to measure motivation suggest that the use of survey items based on the specific requirements of the learning task is a promising approach (Skulmowski & Rey, 2020). The failure to properly measure the construct of value may justify an approach, as seen in Hwang et al. (2013), that includes measures of both usefulness and ease of use. This approach shifts the focus of investigations of educational technology from the value ascribed to it by students to their likelihood to accept it as a valid means of learning. Such a shift suggests value is found in how technology may provide more effective leveraging of teacher expertise (Boettcher, 2007).

In regard to cognitive load, the experimental design should have more strictly controlled the means of content interaction between experimental groups (Skulmowski & Rey, 2019). It is possible that a transient information effect may have been caused by usability features and not features of AR technology. In the AR version of the task, video scrubbing was not possible in the HP Reveal app. Instead, once activated, videos played on a loop that could only be interrupted by moving a device's camera away from a marker. This is not a feature of AR, but of the app used to deliver the AR content. This is in contrast to the videos in the QR code version, which were hosted on YouTube and could be played, paused, and scrubbed through – moving back or forth in the video timeline – to quickly get to particular points in the learning content that may have been difficult. The lack of scrubbing may have made using the AR technology more cognitively demanding for students. Therefore, efforts to determine the source of differences between AR and QR codes in the current study were undermined by tool-specific constraints and not generalizable features of AR technology. This highlights the need for rigorous testing of tools to avoid such constraints problematizing research results (Chwo et al., 2018).

As collaboration was shown to be an important aspect of AR-enabled location-based learning (Morrison et al., 2009; Perry, 2021), the current research could have been improved by properly defining the construct and providing a suitable tool for its measurement. A combination of student perceptions and other means of investigating interactions, such as observations of student behavior, including interaction patterns and content-processing behavior, would have added insight regarding the connection between collaborative learning and motivation, value, and difficulty (Perry, 2021). Knowledge of how much or little information students wrote down regarding directions, for example, may have greatly assisted the discussion of task difficulty.

Conclusion

Location-based learning was facilitated with both AR and QR codes through the cognitive dimension of value, the affective dimension of motivation, and the social dimension of collaboration, which was very highly perceived (Lee & Park, 2019). What, then, was the effect on the student experience of AR technology use over QR code use? It seems that the features of AR that make it suitable for learning are also what may potentially make AR-enabled location-based tasks more difficult. AR made the task more motivating by making the learning content more fun to watch, helping engage students in challenging learning experiences (Pombo et al., 2017). Also, AR made working with classmates more possible through 'place-setting' (Morrison et al., 2009), a major strength of its use in educational settings (Dunleavy & Dede, 2013) which may make a task less demanding. Therefore, AR should be used mainly with pedagogical approaches that prioritize the social construction of knowledge (Morrison et al., 2009) and involve requirements for production, which seem to enhance the negotiation of meaning (Perry, 2021). To fully realize the potential of AR technology to deliver meaningful learning it need not necessarily prioritize location-based learning, as long as AR features are deployed with a view to how they facilitate pedagogy in interactions with the real physical environment (Morrison et al., 2009). Whether that environment be campus locations or portable materials used in collaborative learning activities, the use of AR, and more broadly, MALL, has to enhance the nature of interactions between teachers, students, and pedagogy (An & Oliver, 2020).

The drawbacks of AR included the making of the smartphone technology more difficult to use, the learning content to be more taxing, and perceived performance to be lower. These findings contradict the majority of studies on the cognitive load of AR, that report lower or equal cognitive load with higher performance for AR technology in comparison to other modes of delivery (Buchner et al., 2022). The potential for AR to increase task complexity (Dunleavy & Dede, 2013) means that learner interaction with learning content should be carefully designed to reduce cognitive load. This reflects a key concern of the design of mobile learning (Çakmak, 2019), meaning that the insights gained in the current study regarding this issue can be applied more generally to the design of the mobile interface in MALL. It has been shown how learning in MALL environments can be made more efficient using content that is succinct, labeled, and chunked (Boettcher, 2007), with how content should be handled explained to learners before it is encountered (Boettcher, 2007; Mayer & Moreno, 2003; Wong et al., 2012). Furthermore, this research can be seen as part of the process of grounding technology in pedagogy through understanding its strengths and weaknesses in regard to educational uses (Chwo et al., 2018). Such a grounding recognizes that it is pedagogy, rather than technology, that affects the learning of language (Salmon & Nyhan, 2013) and occurs through detailed planning. Further planning stages include preparing for a MALL system to be used for at least five years, examining how technology use will fit in the existing curriculum, integrating the MALL into the syllabus and grading and providing formative and summative feedback, and the cyclical analysis of MALL use. Such planning leads to opportunities for research with greater validity that contributes to successful long-term curriculum integration of MALL technologies (Chwo et al., 2018), ensuring that the educational benefits of MALL are based on both understanding the potential of a technology and how it can be integrated into a learning environment (Perry, 2021). It is hoped that this research contributes to understanding both of these dimensions and can inform future technology use in educational environments that lead to meaningful learning experiences.

References

An, T. & Oliver, M. (2020). What in the world is educational technology? Rethinking the field from the perspective of the philosophy of technology. *Learning, Media and Technology, 46*(1), 6–19. https://doi.org/10.1080/17439884.2020.1810066

Azuma, R. T. (1997). A survey of augmented reality. *Presence: Teleoperators and Virtual Environments, 6*(4), 355–385.

Boettcher, J. V. (2007). Ten core principles for designing effective learning environments: Insights from brain research and pedagogical theory. *Innovate: Journal of Online Education, 3*(3).

Buchner, J. & Kerres, M. (2021). The impact of augmented reality on cognitive load and performance: A systematic review. *Journal of Computer Assisted Learning, 38,* 285–303. https://doi.org/10.1111/jcal.12617

Çakmak, F. (2019). Mobile learning and mobile assisted language learning in focus. *Language and Technology, 1*(1), 30–48.

Chwo, G. S. M., Marek, M. W., & Wu, W. C. V. (2018). Meta-analysis of MALL research and design. *System, 74,* 62–72. https://doi.org/10.1016/j.system.2018.02.009

Collins, A. (1988). *Cognitive apprenticeship and instructional technology* (Technical Report 6899): BBN Labs Inc.

Conceicao, S. C. O. & Howles, L. L. (2021). *Designing the online learning experience: evidence-based principles and strategies*. Routledge. https://doi.org/10.4324/9781003444121

Council of Europe (2001). Common European framework of reference of languages: Learning, teaching, assessment. Retrieved from: https://www.coe.int/en/web/common-european-framework-reference-languages (accessed June 11th, 2019)
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.
Dunleavy, M., & Dede, C. (2013). Augmented reality teaching and learning. In J. M. Spector et al. (eds.), *Handbook of research on educational communications and technology,* 735–745. https://doi.org/10.1007/978-1-4614-3185-5_59
Guariento, W., & Morley, J. (2001). Text and task authenticity in the EFL classroom. *ELT Journal, 55*(4), 347–353. https://doi.org/10.1093/elt/55.4.347
Hart, S. G. (2006). Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 50*(9), 904–908. https://doi.org/10.1177/154193120605000909
Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Human Mental Workload,* 139–183. https://doi.org/10.1016/s0166-4115(08)62386-9
Herrington, J. & Oliver, R. (2000). An instructional design framework for authentic learning environments. *Educational Technology Research and Development, 48*(3), 23–48.
Hockly, H. (2019). Technology for the language teacher: augmented reality. *ELT Journal, 73*(3), 328–334. https://doi.org/10.1093/elt/ccz020
Holden, L. C., & Sykes, J. M. (2011). Leveraging mobile games for place-based language learning. *International Journal of Game-Based Learning, 1*(2), 1–18.
Hoopingarner, D. (2009). Best practices in technology and language teaching. *Language and Linguistics Compass, 3*(1), 222–235. https://doi.org/10.1111/j.1749-818x.2008.00123.x
Howard, M. C. (2016). A review of exploratory factor analysis decisions and overview of current practices: What we are doing and how can we improve? *International Journal of Human-Computer Interaction, 32*(1), 51–62. https://doi.org/10.1080/10447318.2015.1087664
Hwang, G.-J., Yang, L.-H., & Wang, S.-Y. (2013). A concept map-embedded educational computer game for improving students' learning performance in natural science courses. *Computers & Education, 69,* 121–130. https://doi.org/10.1016/j.compedu.2013.07
Kan, T. W., Teng, C. H., Chen, M. Y. (2011). QR code based augmented reality applications. In, Furht, B. (eds) *Handbook of augmented reality.* Springer. https://doi.org/10.1007/978-1-4614-0064-6_16
Law, C., & So, S. (2010). QR codes in education. *Journal of Educational Technology Development and Exchange, 3*(1), 85–100.

Lai, A.-F., Chen, C.-H., & Lee, G.-Y. (2019). An augmented reality-based learning approach to enhancing students' science reading performances from the perspective of the cognitive load theory. *British Journal of Educational Technology, 50*(1), 232–247. https://doi.org/10.1111/bjet.12716

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* Cambridge University Press.

Lee, S-M. & Park, M. (2020). Reconceptualization of the context in language learning with a location-based AR app. *Computer Assisted Language Learning, 33*(8), 936–959. https://doi.org/10.1080/09588221.2019.1602545

Liu, T. Y., Tan, T. H., & Chu, Y. L. (2010). QR code and augmented reality-supported mobile English learning system. *Mobile multimedia processing: Fundamentals, methods, and applications,* 37–52.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 1*(38), 43–52.

McGuiggan, S., Kosturko, L., McGuiggan, J., and Sabourin, J. (2015). *Mobile learning: A handbook for developers, educators, and learners.* Wiley.

Morrison, A., Oulasvirta, A., Peltonen, P., Lemmelä, S., Jacucci, G., Reitmayr, G., Näsänen, J., & Juustila, A. (2009). Like bees around the hive: A comparative study of a mobile augmented reality map. *CHI '09: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems 2009,* 1889–1898.

Parmaxi, A. & Demetriou, A. A. (2020). Augmented reality in language learning: A state-of-the-art review of 2014–2019. *Journal of Computer Assisted Learning.* https://doi.org/10.1111/jcal.124861–15

Perry, B. (2021). Gamified mobile collaborative location-based language learning. *Frontiers in Education, 6,* 1–15. https://doi.org/10.3389/feduc.2021.689599

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning, 64*(4), 878–912. https://doi.org/10.1111/lang.12079

Pombo, L., Marques, M. M., Lucas, M., Carlos, V., Loureiro, M, J., & Guerra, C. (2017). Moving learning into a smart urban park: students' perceptions of the Augmented Reality EduPARK mobile game. *Interaction Design and Architecture(s) Journal, 35,* 117–134.

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

Reinders, H., Lakarnchua, O., & Pegrum, M. (2015). A trade-off in learning – Mobile augmented reality for language learning. In M. Thomas & H. Reinders (Eds.), *Contemporary task-based language teaching in Asia,* 244–256. Bloomsbury.

Rikala, J. & Kankaanranta, M. (2012). The use of quick response codes in the classroom. International Conference on Mobile and Contextual Learning.

Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology, 43*(3), 450–461.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*(1), 54–67.

Salmon, J. & Nyhan, J. (2013). Augmented reality potential and hype: towards an evaluative framework in foreign language teaching. *Journal of Language Teaching and Learning, 1,* 54–68.

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

Skulmowski, A. & Rey, G. D. (2020). Subjective cognitive load surveys lead to divergent results for interactive learning media. *Human Behavior & Emerging Technologies, 2,* 149–157. https://doi.org/10.1002/hbe2.184

Squire, K. D. (2009). Mobile media learning: multiplicities of place. *Horizon, 17*(1), 70–80. https://doi.org/10.1108/10748120910936162

Sun, Y. & Gao, F. (2020). An investigation of the influence of intrinsic motivation on students' intention to use mobile devices in language learning. *Educational Technology Research Development, 68,* 1181–1198. https://doi.org/10.1007/s11423-019-09733-9

Sweller, J. (2011). Cognitive load theory. *Psychology of Learning and Motivation, 55,* 37–76. https://doi.org/10.1016/b978-0-12-387691-1.00002-8

Taylor, S. & Stone, A. (2018). Smartphone augmented reality for EFL learners. *2018 PanSIG Journal,* 258–263.

Taylor, S. & Stone, A. (2023). Ready, set...? End-user testing of an AR treasure hunt learning game. *Playful CALL 2022: Selected papers from the JALTCALL 2022 Conference, Japan,* 6884. https://doi.org/10.37546/JALTSIG.CALL.PCP2022-04

Wong, A., Leahy, W., Marcus, N., & Sweller, J. (2012). Cognitive load theory, the transient information effect and e-learning. *Learning and Instruction, 22*(6), 449–457. https://doi.org/10.1016/j.learninstruc.2012.05.004