applied sciences

MDPI

Article

Multi-Modalities in Mobile Technology for Assisted Learning Performance in Higher Education in China

Ruichen Yuan *, Habibah Ab Jalil and Muhd Khaizer Omar

Faculty of Educational Studies, Universiti Putra Malaysia, Serdang 43400, Selangor, Malaysia; habibahjalil@upm.edu.my (H.A.J.); khaizer@upm.edu.my (M.K.O.)
* Correspondence: gs60475@student.upm.edu.my

Abstract: Mobile technology, especially mobile learning, has long been an emerging and thriving field, and remains a main theme in mobile learning applications and systems. The extensive utilization of mobile learning has prompted the invention of many mobile applications. As a result of rapid advances in application technologies, various learning applications can combine different media or multi-modalities, such as video, audio, images, animated graphics, and text, to create multimedia learning resources that engage learners. However, the most favorable modalities in different learning applications that assist performance are worth exploring. This study employed mixed methods to investigate the current multi-modality situation in learning application utilization among 300 university students in China, where a rapid educational technology revolution is occurring. The findings revealed that the verbal modality (M = 3.99, S*D = 0.79) and the writing modality (M = 3.99, S*D = 0.75) in the learning applications were less enjoyable and less effective at enhancing learning performance. In exam-based or function-based apps, all five modalities in this research were considered important, especially the visual and aural modes. The results of this study also revealed that a majority of university learners were satisfied with the multi-modalities in different types of applications, except for game-based apps, that assist their learning performance (56.7%, M = 3.87, S*D = 0.79), which contrasts with the results of several related studies. Overall, college users perceived that multi-modalities were effective in helping them to complete tasks, and all modalities in current applications satisfied most of the users' needs to assist their learning performance. In the end, the findings indicated a positive and strong linear relationship [r = 0.766, $p < 0.05$] between multi-modalities and assisted learning performance with the help of more capable (knowledgeable) others with the use of mobile applications.

Keywords: assisted learning performance; multi-modalities; mobile technology; mobile learning applications; higher education institutions

Academic Editors: Douglas O'Shaughnessy and Antonio Pagliaro

Received: 10 February 2025
Revised: 3 March 2025
Accepted: 7 March 2025
Published: 10 March 2025

Citation: Yuan, R.; Ab Jalil, H.; Omar, M.K. Multi-Modalities in Mobile Technology for Assisted Learning Performance in Higher Education in China. *Appl. Sci.* **2025**, *15*, 2987. https://doi.org/10.3390/app15062987

1. Introduction

Recent years have seen the unprecedented proliferation of mobile devices in individuals' daily lives, and the use of mobile technology to support education continues to grow. It was observed that today's students are equipped with various mobile devices that connect them to the world regardless of time and place, allowing them to access information wherever and whenever they want [1]. Mobile devices were defined as technological tools that meet certain user needs, particularly the need to enhance education purpose and knowledge acquisition [2]. Mobile technology has become a significant tool in education owing to its ease of use and versatility. For many educators, mobile technology in the field of teaching and learning has recently become an important area of research [3].

As mobile devices proliferate, their use and potential in higher education have become common topics in educational research and practice. The new generation of mobile devices enables students to learn and collaborate anytime and anywhere, and has become a significant component of higher education [4]. Today's generation of students is familiar with multiple types of mobile technology [5–7]. It was predicted that the ubiquity, personalization, and socialization of mobile technology over the next 10 years, with the potential for transforming the nature of teaching and learning [8].

Therefore, mobile learning remains an active area of research, requiring attention from diverse perspectives. The term mobile learning, or m-learning, is understood as a process of learning through mobile phones [9]. It was stated that, at present, mobile learning apps represent the latest tool for demonstrating learning materials that are designed to be self-paced, holistic, interactive, and visual, providing contextualized input and offering various options for enhancing a learner's skills [10]. As the capabilities of mobile technology continue to grow and become more ubiquitous, mobile applications are in high demand around the world. Institutions of higher education have shown significant interest in the use of technological devices in education, and numerous mobile phone applications have been increasingly employed in teaching and learning [11].

Mobile learning features can help students become more self-motivated and self-directed learners. Mobile learning also has the potential to improve engagement and assist the learning process, as well as increase efficient learning for students [12]. Mobile learning is the use of portable devices such as smartphones, iPads, and tablets to provide anywhere and anytime learning experiences [13–15]. Therefore, mobile learning applications have the potential to provide broader access to teaching and learning resources, which may improve learner performance and outcomes. Institutions of higher education and contemporary educators have an array of digital technologies at their disposal for enhancing teaching and learning process and outcomes. It was observed that college students are bringing mobile devices like smartphones and tablets to educational settings, likely perceiving that these technologies are significant to their academic success [16].

Therefore, it is worthwhile to investigate how to better integrate multi-modal features into mobile learning applications to more effectively support student engagement and learning. It was indicated that the speedy progress of app technologies has allowed them to combine different media, known as multi-modality [17]. For instance, videos, audio, pictures, animation, and texts are combined to create multimedia instructional materials, which stimulate students' interest in learning. The rampant use of technology has empowered this generation, enabling it to produce and not merely consume content. Such content is not limited to text and involves multiple media (sound, images, videos, animations) [18]. In their systematic review, however, the authors emphasized the limitations of mobile learning applications in fostering student engagement, and the need for further research on multi-modal features that can enhance learning outcomes [19]. Therefore, the kinds of media and modalities that are useful in different learning applications are worth exploring.

2. Mobile Technology for Learning

Mobile technologies attract new users at an unpredictable speed, offering increased capacity and more sophisticated functions in many fields. It was pointed out that mobile technology is an important component of empirical research that illustrates how learning with these devices creates opportunities to directly influence learning processes or outcomes, as well as opportunities to collect previously unavailable data to better understand and model the learning process [20]. The benefits of mobile technology for learning and teaching have been widely cited by researchers. For example, it was emphasized that mobile technology can be used to facilitate learning, communicate with colleagues, explore

different perspectives, obtain information, and access learning materials [21]. In facilitating improved individual learning performance, mobile technology plays a critical role. According to [22], an extensive body of research addresses the critical role of mobile technology in the effectiveness and efficiency of learning. It was suggested that mobile technologies have great potential for language learning among college students [23].

3. Mobile Learning Applications

In the 21st century, people are familiar with the concept of applications. App is the short form of the phrase "application software", which is generally downloaded from "app stores such as App Store, Google Play, Windows Phone Store, and BlackBerry App World" [17]. However, in China, people typically use the App Store or Android Store to install apps for entertainment, gaming, education, or daily needs. The rapid advancement of mobile technology, especially mobile applications that can assist learning, has attracted increasing interest from educators. Undeniably, the research on mobile learning applications is thriving and in its prosperous phase. Because of these emergent technologies, the use of mobile devices as learning tools is now being fully explored and remains an area of active research [24]. It was also indicated that, as a projection for future research, the current use of mobile learning applications can be further explored by focusing on users' opinions of specific applications for specific needs [25]. It was found that each app offers a unique way of learning, and, obviously, each has its pros and cons. No one app is the best one, which implies that a combination of apps makes learning more efficient [26]. More importantly, exploring which features or modes are most appreciated by college users when they employ learning applications is a promising research direction.

4. Assisted Learning Performance in the Zone of Proximal Development (ZPD)

The Zone of Proximal Development theory proposes that the instructor or teacher is in the external sphere as a facilitator to guide the learner when the required knowledge is beyond the learner's comprehension. Coincidentally, [27] also discussed the necessity for teachers to effectively supervise and guide students in their mobile learning processes. For students, achieving better learning performance commonly requires assistance, mainly from teachers and more able peers, which is based on Zone of Proximal Development (ZPD) from [28]. In the study of [29], more capable others include parents, teachers, experts, peers, etc., and, in this research, the assistance is mainly provided by teachers and more capable peers.

Learners who use mobile learning technology/applications to achieve their goals need guidance and assistance from more capable (knowledgeable) others, like teachers or instructors and even outstanding peers. Teaching occurs when assistance is offered at points in the Zone of Proximal Development (ZPD) where performance requires assistance [29]. The Zone of Proximal Development (ZPD) displays learners' and teachers' roles in the learning process. The authors, therefore, stressed that teaching must be redefined as "assisted performance" in that it occurs when performance is achieved with assistance [30]. Therefore, in this study, assisted learning performance mainly refers to learners obtaining support or guidance from their teachers or more capable peers using mobile technology (applications) in higher education.

5. Multi-Modalities in Learning Application Research in China's Higher Institutions

In recent years, the exploration of multi-modalities in learning applications has gained prominence in educational research, particularly within the context of China. This trend

reflects a broader understanding of how diverse learning modalities can enhance educational outcomes and cater to the varied needs of learners. Multi-modal learning, which integrates various forms of content delivery—such as visual, auditory, and kinesthetic—has been shown to significantly improve engagement and retention among students [31]. For instance, they highlight that multi-modal learning strategies in medical education can guide deep learning-based segmentation networks to extract complementary information from different modalities, thereby achieving superior performance compared to single-modal approaches [31]. Furthermore, the rise of multi-modal machine learning techniques has opened new avenues for research and application in educational settings. The study of [32] provide a comprehensive survey of multi-modal machine learning, discussing its potential applications in educational technology and the importance of coordinated representations for effective learning outcomes. According to [33], they explored deep learning-based image segmentation in multi-modal medical imaging, reflecting the growing trend of integrating advanced technologies in educational applications, particularly in fields like medical education and training. Different learning styles among medical students were examined and the significance of multi-modal learning strategies was also highlighted [34]. These studies indicate that students who engage with multiple sensory modalities tend to perform better, suggesting that multi-modal approaches can enhance learning outcomes in higher education.

6. Purpose of the Study

Learning apps available in the app stores are numerous, and they offer a variety of multimedia modes and functions, such as sounds, videos, music, or images, for personal, perceptual, and field-independent learning [35]. An abundance of applications enables communication, image, video, and audio capture and sharing in multiple settings [36]. These multi-modalities in mobile learning apps allow college students in various disciplines or major subjects to benefit from versatile applications to enhance their learning performance and ease the complex learning process. Therefore, to maximize the benefits of smartphone use, educators must carefully adapt the technology to specific learning goals and outcomes to ensure that meaningful learning takes place [4]. It was observed that, with a plethora of applications allowing users to communicate, capture, and share images, video, and audio in various settings, educators and students can benefit from easy and frequent access [36]. However, with the rapid development of mobile apps, the implications of utilizing specific functions of learning applications are worth exploring specifically. This allows us to evaluate which modes are most interactive and valuable for apps aimed at various subjects or majors. Therefore, new perspectives on the multi-modality of the specific features of various mobile learning applications were explored.

This research study is part of a larger project that investigates the adoption strategies of mobile technology for assisted learning performance among students in dozens of Chinese universities. Students' perceptions of mobile technology, particularly the assistance they receive from users who are more knowledgeable about mobile learning applications and the benefits of multi-modality in applications to their learning performance, are investigated through questionnaires and surveys. This paper reports data from 300 students, mainly on their perspectives of multi-modalities in learning applications. The quantitative and qualitative findings were combined to interpret the results.

7. Methods

7.1. Aim and Research Questions

To contribute new and more precise data to the existing research, the focus of this study is exploring students' perspectives on their assisted learning performance when using

mobile technology, especially with the assistance they receive from more knowledgeable users when employing various mobile learning applications and the benefits of multi-modality in learning applications to their learning performance.

The following research questions guide this study:

Research question 1: What are learners' perspectives on assisted learning performance from multi-modalities in mobile learning applications?

Research question 2: What are learners' perspectives on the most useful modalities of mobile learning apps for assisting their learning performance?

Research question 3: What is the learners' perceived satisfaction of multi-modalities in different mobile learning applications?

Research question 4: What is the relationship between multi-modalities and learners' assisted learning performance from more capable others (their teacher or peers) with the use of mobile technology (applications)?

7.2. Participants and Universities

In this study, 200 respondents were recruited from various disciplines and year levels, i.e., freshmen, sophomores, juniors, and seniors, from various Chinese universities, mainly from Xi'an, Shaanxi Province, where many universities afforded a diversity of the respondents. The remaining 100 participants studied at a private college located in Xi'an. They were undergraduates in the department of humanities and education who were randomly selected. The selection criteria were, firstly, they must be a college student in any major and, secondly, they must have more than two years' mobile learning experience.

Of the 300 Chinese college students in this study, 42.33% were males and 57.67% were females, indicating a higher percentage of female participants, with a mean score of 1.58. First-year students accounted for most of the respondents at 139 (46.3%), followed by juniors at 78 (26%), and senior students at 69 (23%). The major subjects in Chinese universities are commonly categorized into 12 types. The highest proportion of respondents majored in education (35.7%) because 100 participants were the result of purposive sampling and were from the education department of a private university. Beyond this group, participants who majored in engineering (16%), literature (13%), management (11.7%), and economics (7.7%) altogether made up the second-largest proportion. As for the location, the first- and second-largest samples of students came from the central (29.3%) and northwest parts (25.3%) of China, as the researcher was located in the northwest part of China. Hence, more resources could be obtained in this area, which meant, to some extent, the sample size was representative of the whole population. The descriptive summary of student profiles is shown in Table 1.

Table 1. Descriptive summary of student profiles.

Variable	Category	Frequency (f)	Percent (%)
Gender	Male	127	42.3
	Female	173	57.7
Grade	Freshman	139	46.3
	Sophomore	14	4.7
	Junior	78	26.0
	Senior	69	23.0

Table 1. *Cont.*

Variable	Category	Frequency (f)	Percent (%)
Major	Philosophy	14	4.7
	Economics	23	7.7
	Education	107	35.7
	Law	4	1.3
	Literature	39	13
	History	1	0.3
	Science	7	2.3
	Engineering Agriculture	48	16
	Medicine	5	1.7
	Management	2	0.7
	Arts	35	11.7
		15	5
Location	Eastern part of China	8	2.7
	Western part of China	33	11
	Central part of China	88	29.3
	Northern part of China	20	6.7
	Southern part of China	52	17.3
	Northeast China Southeast	1	0.3
	China Northwest China	9	3
	Southwest China	76	25.3
		13	4.3

7.3. Instruments

This study employed a mixed-method design to approach the research questions from a holistic perspective. Using only one research method may not be sufficient to address research issues or formulate research questions. Moreover, the quantitative and qualitative methods complement one another, enabling more robust analyses. Using a mixed-method design allows the researcher to benefit more from using both research methods.

The first tool used for collecting quantitative data was the questionnaire. It was noted that online surveys have been used extensively as data collection tools in research due to their efficiency and flexibility [37]. The investigator needed to gather information on overall perspectives on multi-modality in mobile applications to assist learning performance among the participants. Therefore, this first data collection method utilized online questionnaires with closed-ended questions to collect information about some pertinent questions related to our research aim. Some questions were developed by [38] and adapted for this study, some were adapted from the Post-Study Usability Questionnaire (version 3), and others were developed especially for this study.

The second main instrument for the collection of qualitative data was the semi-structured interview. According to [39], the semi-structured interview provides reliable and comparable qualitative data and allows informants the freedom to express their views on their own terms. Semi-structured interviews were conducted one-on-one and lasted from thirty minutes to one hour. The aims and a simple introduction were conveyed to the students at the beginning of the interview, and participants had to use certain apps on their mobile phones to answer some research questions. Other than that, they did not need to prepare anything beforehand. The main interview questions are provided in Table 2.

Table 2. Semi-structured interview questions.

1	Would you prefer to search for a learning app on your own or ask others when you need one? And why? If on your own: What do you usually use to search for on your own? If ask others: Who do you often turn to for assistance?
2	What do you think of the importance of obtaining some recommendations from teachers or peers when choosing suitable learning apps during the whole university study process?
3	Do you think their recommendations or assistance to choose and use the suitable apps can help you improve your academic performance? And why?
4	What type of learning app do you use most often? And why?
5	Do you think the use of learning apps can improve your learning performance? And why?
6	Is there any function that you need but do not provide in the learning apps?
7	Among multi-modalities, such as pictures, text, audio, video and others, which modality do you think is more useful in your often used learning apps?
8	Do you think which category of app needs visual modality (pictures), aural modality (audio), verbal modality (sound recording), reading modality (text) and writing modality (word typing) most, respectively? And why?
9	What do you think of the significance of having these modalities to improve your learning performance?
10	Is there any modality that you need but do not have in a certain app so far?

7.4. Data Collection and Analysis

To provide data, 300 participants took part in the online questionnaire. The survey consisted of 62 questions and was divided into four sections, which took around 10–15 min to complete. The first section included questions that related to perspectives on assisted learning performance through more capable (knowledgeable) others using mobile learning applications. The second section contained questions that covered perspectives on assisted learning performance through multi-modalities in mobile learning applications. The third section included questions that represented perspectives on five modalities that assist learning performance in different learning applications. The fourth section of the survey consisted of questions that represented participants' perspectives on their perceived satisfaction with multi-modalities in mobile learning applications. In this quantitative part, the collected data were analyzed using the Statistical Package for Social Science (SPSS), version 26. First, normal distribution was assessed using the Kolmogorov–Smirnov test. Then, a reliability test was conducted representing the perspectives, which was checked by calculating Cronbach's alpha. Cronbach's alpha reported a range of between 0.913 and 0.983, confirming that the questionnaire was reliable as it showed good internal consistency. To examine the fourth research question determining the strength of the relationship between the variables of study, a Pearson product–moment correlation was employed.

To collect data for the semi-structured interviews, 30 volunteers were selected randomly from the 300 respondents to explore their perspectives on their assisted learning performance using mobile technology, especially regarding the assistance from more knowledgeable users when employing various mobile learning applications and the assistance from multi-modality. A diverse group of 30 respondents, identified and assigned numbers 1 to 30, were interviewed. The participants, 13 males and 17 females, were university students in different major subjects, years, and universities in China. To provide a more

comfortable and relaxed environment for interviewers to elicit candid answers, the interviews were held one-on-one in Chinese. They were held online using apps like Zoom or Tencent Meeting, and each one lasted from thirty minutes to one hour. After the interviews, the responses were collected, translated, and analyzed to provide solid responses to the research questions.

When transcribing interviews, Microsoft Word was used to manage data. An efficient way to retrieve the entire oral production of data is to transcribe it fully, which aids the researcher in examining and analyzing the transcript repeatedly and effectively. Transcription in qualitative research helps the researcher to organize and analyze data systematically [40]. The researcher tried to avoid alterations and misinterpretations in transcribing data from the semi-structured interviews. As for coding, all the data were analyzed with Nvivo, which was relatively easy to use. It is possible to import documents directly from a word processing package and code them easily on the screen, followed by sessions of carefully reading the transcriptions to determine what phrases may be considered "strategy-like" expressions. Coding stripes can be made in the margins of documents so that the researcher can see, at a glance, which codes have been used where [41]. Further, the researcher also employed selective coding, "the process of selecting the core category, systematically relating it to other categories, validating these relationships and filling in categories that need further refinement and development" [42].

Regarding theme construction, it is important to note that the audience played a vital role. The data were first descriptively labeled as "units of meaning" [43] and initially coded at a relatively lower conceptual level. By adopting a grounded approach to coding guided by the audience's insights, the emerging codes steered the rest of the data analysis process. The initial codes were then grouped into more abstract themes. After the review, definition, and naming of the themes, the initial results were examined, and the expected strategies were deduced as expected, with the audience's contributions in mind.

In the end, the quantitative results and qualitative results were contrasted and compared according to each research question so that the final interpretation was generated holistically.

8. Results

8.1. Qualitative Results: Assisted Learning Performance Through Multi-Modalities

It is believed that multi-modalities in learning applications can improve the entire learning experience, making the learning process more entertaining and relaxing. Therefore, the learning outcomes improve, and the learning performance is also better assisted. Interviewee 10 commented the following:

> *It is more vivid and easier to understand the content. It is more attractive, makes people more interested in learning, and the learning performance will be better.*

Interviewee 12 gave a similar and detailed response:

> *I think each modality has its own unique effect. For example, the visual modality will impact our cerebral cortex and make it easier to remember some complicated things. The specific operations and functions can be understood through pictures. I think auditory modality and speech modality will also impact our brain cells and enhance our memory of knowledge, while the combination of multiple modalities will more significantly improve our learning effects and performance.*

The second discovery from the analysis revealed that the visual modality is more appreciated by users in their preferred learning apps for many reasons, such as the vivid, attractive, colorful, and straightforward effects shown in Figure 1.

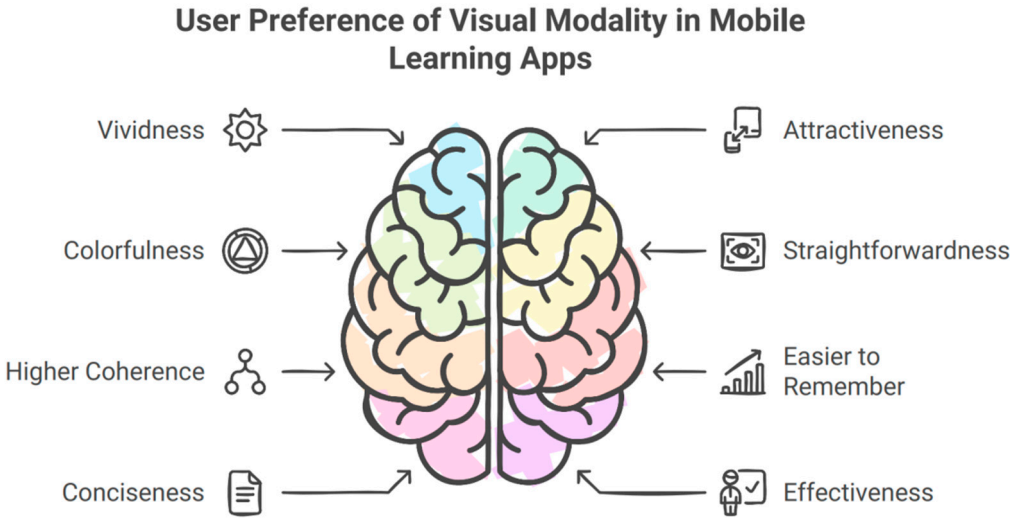


Figure 1. User preference of visual modality in mobile learning apps.

> *I perceive that videos are more useful; Interesting videos are more attractive than pictures or boring text or audio. Videos will make me continue watching and selectively remember the knowledge points that I am interested in or need.*
>
> (Interviewee 2)

Interviewee 8 made a similar justification:

> *I think videos are the most useful one. Compared with pictures, videos have higher coherence, connect you with knowledge, are vivid and colorful, and are easier to remember; videos are more straightforward, objective, vivid and concise than text; videos are more effective than audio.*

However, some learners provided more holistic answers, like Interviewee 19:

> *Which modality is more useful depends on the learning content, goals, and personal learning style and preferences. For content that is highly theoretical and requires in-depth thinking, text may be more effective; for content that requires visual understanding or operational demonstration, pictures and videos may be more suitable; in mobile learning or busy situations, audio may be a good and convenient choice.*

The third finding from the interview content on this theme is related to the application categories created by the university students themselves according to their own learning experience. The answers are diverse but enlightening. The underlying logic of the classification is based on the aim and the function of the learning applications, as Interviewee 6 replied:

> *There have been classifications, and the categories include vocabulary inquiry, reading, English speaking practice, and course viewing. The classification criteria are the main functions of the app for me, or what I want to get from the app.*

Some learners distinguished more application types based on their daily study, such as subject-based, function-based, exam-based, and task-based apps. We learned this from the responses of Interviewees 12 and 17:

> *I have classified the learning apps on my mobile phone. They were divided according to the type of learning and the subjects. For example, according to the type of learning, they were divided into English and other subjects. Then English can be divided into individual category due to different functions, like grammar, composition, etc.; if divided according to the type of examination, there were categories like computer level 2 examination, teacher qualification certificate examination, CET 4 and CET 6 examination, etc.*

Yes, they were classified. Generally speaking, they are classified according to specific learning tasks, such as video courses, online Q&A, question discussions, and question solving software, etc.

Comparing the learners' responses, we also discovered that many apps pertain to diverse subjects or majors, functions, tasks, or exams, or are game-based apps like Duolingo. The classification is shown in Figure 2.

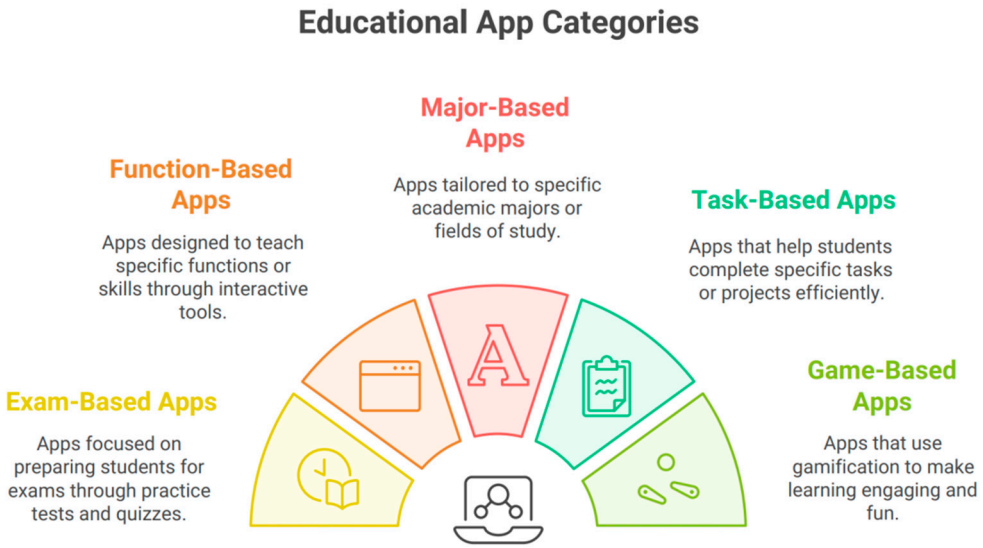


Figure 2. Educational app categories.

The fourth discovery is associated with the modality that is most useful or suitable for particular learning apps, as demonstrated in Figure 3. Because classifying learning apps is difficult and multifaceted, the responses in this part were heterogeneous. Still, some common points stand out. Of the numerous varieties of learning apps, language learning apps or English exam apps appeared repeatedly in the college students' comments. They believed that, in these exam- and function-based apps, all five modalities are important, especially the visual mode and aural mode, as in Interviewee 17's response:

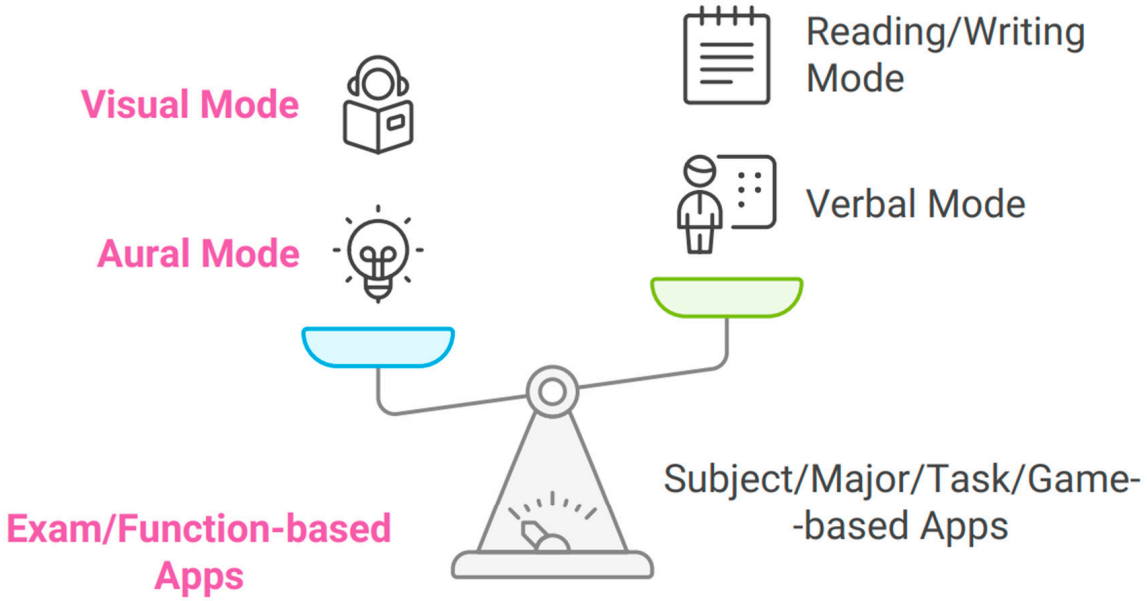


Figure 3. Modality importance in different types of learning apps.

Function-based apps, such as word memorization apps, dictionary apps, and video course apps, I think they require more visual and aural modalities. For exam apps like the CET-4 and CET-6 exams, the aural and reading modes are more important. For apps with spoken language functions, the verbal function is even more important. It really depends.

The detailed analysis of these data revealed that most university learners believe that multi-modalities in mobile learning applications have a positive influence on their learning performance, owing to their many ways of displaying information and knowledge. From the five modalities, the visual mode, such as videos and pictures, was most preferred by college learners for its entertaining and visual impact, improving memory retention. In terms of classifying the learning applications, although the answers were multifarious, commonalities could still be identified. The categories were subject-based, major-based, function-based, task-based, exam-based, and game-based apps, and, among them, game-based apps were the least mentioned.

8.2. Quantitative Results

8.2.1. Assisted Learning Performance Through More Capable Others

The findings concerning university learners' responses to each item represent the dimension "Assisted Learning Performance through More Capable Others", as shown in Table 3. More than 80% of university participants agreed and strongly agreed with the seven items under this dimension. The item "Experts or more capable others (my teacher or other peers) always instruct me to use his/her recommended mobile learning applications in class or after class" had the most disagree and strongly disagree responses among the seven items (M = 3.88, S*D = 0.81). The statement with the highest mean score for expert support was "It is important to draw on experts or more capable others (my teacher or other peers) for their suggestions or guidance when actually using a mobile learning application" (M = 4.02, S*D = 0.75), and the item "Experts or more capable others (my teacher or other peers) always instruct me to use his/her recommended mobile learning applications in class or after class" also scored the lowest on the scale (M = 3.88, S*D = 0.81). The overall mean score for this scale was 3.97, and the standard deviation was 0.76. The results indicated the importance of having guidance and assistance from more experienced individuals like experts, teachers, and peers when using mobile technologies (applications). More importantly, a significant portion of students believed that they needed instruction in using mobile learning applications, while there was not enough support given in or out of class.

Table 3. Descriptive statistics for assisted learning performance through more capable others.

Statement	N	Percent (100%)					M	S*D
		(f)	(f)	(f)	(f)	(f)		
		SD	D	LA	A	SA		
I always ask for advice from experts or more capable others (my teacher or other peers) when choosing a mobile learning application.	300	0.3	4	19.7	55.7	20.3	3.92	0.77
		(1)	(12)	(59)	(167)	(61)		
It is important to draw on experts or more capable others (my teacher or other peers) for their suggestions or guidance when choosing a mobile learning application.	300	0.3	4.3	15.7	56.7	23	3.98	0.77
		(1)	(13)	(47)	(170)	(69)		
It is important to draw on experts or more capable others (my teacher or other peers) for their suggestions or guidance when actually using a mobile learning application.	300	0.3	4	12.7	59	24	4.02	0.75
		(1)	(12)	(38)	(177)	(72)		
Experts or more capable others (my teacher or other peers) are a great help to me in choosing a mobile learning application.	300	0.3	4	13	59.7	23	4.01	0.74
		(1)	(12)	(39)	(179)	(69)		

Table 3. *Cont.*

Statement	N	Percent (100%)					M	S*D
		(f)	(f)	(f)	(f)	(f)		
		SD	D	LA	A	SA		
Experts or more capable others (my teacher or other peers) always share useful mobile learning applications in or after class.	300	0.3	4	15.3	57.3	23	3.99	0.76
		(1)	(12)	(46)	(172)	(69)		
Experts or more capable others (my teacher or other peers) always instruct me to use his/her recommended mobile learning applications in class or after class.	300	0.3	6.3	18	55.7	19.7	3.88	0.81
		(1)	(19)	(54)	(167)	(59)		
My academic performance has improved with the assistance of experts or more capable others (my teacher or other peers) with the use of mobile learning applications.	300	0	3.7	17.7	56.7	22	3.97	0.74
		(0)	(11)	(53)	(170)	(66)		
Overall							3.97	0.76

Scale description: SD = strongly disagree, D = disagree, LA = less agree, A = agree, SA = strongly agree; M = mean; S*D = standard deviation.

8.2.2. Assisted Learning Performance Through Multi-Modalities in Mobile Learning Applications

Table 4 is based on a descriptive analysis of the learners' responses to the items related to perspectives on multi-modalities. A majority of the students agreed and strongly agreed that they enjoyed using the five individual modalities and perceived that they were very important in enhancing their learning performance. On the other hand, the same percentage of respondents agreed less with the item "I enjoy using the visual modality (e.g., pictures, images or videos) in learning applications" (9%, M = 4.11, S*D = 0.71) and the item "The aural modality (e.g., audio or listening function) in the learning applications is very important to enhance my learning performance" (9%, M = 4.13, S*D = 0.69). The percentage was also the lowest among the 10 items. However, a slight difference could be seen in the mean score. The highest mean score was on the item where users declared that the aural modality (e.g., audio or listening function) in the learning applications was very important in enhancing their learning performance (M = 4.13, S*D = 0.69), while the lowest mean scores were on the item "I enjoy using the verbal modality (e.g., speaking or sound-recording function) in the learning applications" (M = 3.99, S*D = 0.79) and "The learning applications' writing modality (e.g., word-typing function) is very important to enhance my learning performance" (M = 3.99, S*D = 0.75). Lastly, the overall mean score was 4.04 (S*D = 0.74). The findings indicate that the visual modality was much more favorable than the other modes, and the aural modality was believed to be more important in improving their learning performance. The verbal modality was less favorable than the others, and the writing modality was less significant in enhancing their learning performance compared to the other four modes.

Table 4. Descriptive statistics for perspectives on assisted learning performance through multi-modalities.

Statements	N	Percent (100%)					M	S*D
		(f)	(f)	(f)	(f)	(f)		
		SD	D	LA	A	SA		
I enjoy using the visual modality (e.g., pictures, images or videos) in learning applications.	300	1	1.7	9	62	26.3	4.11	0.71
		(2)	(5)	(27)	(186)	(79)		
The learning applications' visual modality (e.g., pictures, images or videos) is very important to enhance my learning performance.	300	0.7	3	13.3	58.7	24.3	4.03	0.75
		(2)	(9)	(40)	(176)	(73)		
I enjoy using the aural modality (e.g., audio or listening function) in learning applications.	300	0.7	2.7	12	57	27.7	4.08	0.75
		(2)	(8)	(36)	(171)	(83)		

Table 4. *Cont.*

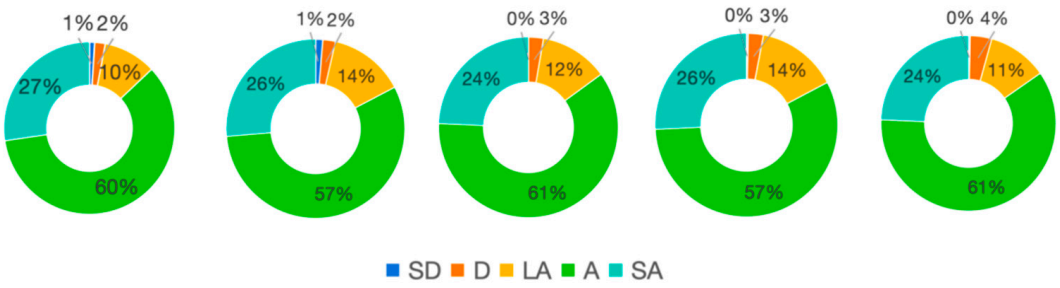
Statements	N	Percent (100%)					M	S*D
		(f)	(f)	(f)	(f)	(f)		
		SD	D	LA	A	SA		
The aural modality (e.g., audio or listening function) in the learning applications is very important to enhance my learning performance.	300	0.3	2.3	9	60.7	27.7	4.13	0.69
		(1)	(7)	(27)	(182)	(83)		
I enjoy using the verbal modality (e.g., speaking or sound-recording function) in the learning applications.	300	0.7	5.3	11.3	59.7	23	3.99	0.79
		(2)	(16)	(34)	(179)	(69)		
The verbal modality (e.g., speaking or sound-recording function) in the learning applications is very important to enhance my learning performance.	300	0	4.3	15.5	56.3	24.3	4.01	0.75
		(0)	(13)	(45)	(169)	(73)		
I enjoy using the reading modality (e.g., text) in the learning applications.	300	0.3	3	15	59	22.7	4.01	0.73
		(1)	(9)	(45)	(177)	(68)		
The reading modality (e.g., text) in the learning applications is very important to enhance my learning performance.	300	0.3	2.7	14.3	59.3	23.3	4.03	0.72
		(1)	(8)	(43)	(178)	(70)		
I enjoy using the writing modality (e.g., word-typing function) in the learning applications.	300	0.7	4.3	11.7	61.3	22	4	0.76
		(2)	(13)	(35)	(184)	(66)		
The learning applications' writing modality (e.g., word-typing function) is very important to enhance my learning performance.	300	0.3	4.3	13.3	59.7	22.3	3.99	0.75
		(1)	(13)	(40)	(179)	(67)		
Overall							4.04	0.74

Scale description: SD = strongly disagree, D = disagree, LA = less agree, A = agree, SA = strongly agree; M = mean; S*D = standard deviation.

8.2.3. Perspectives on the Five Modalities in Different Learning Applications

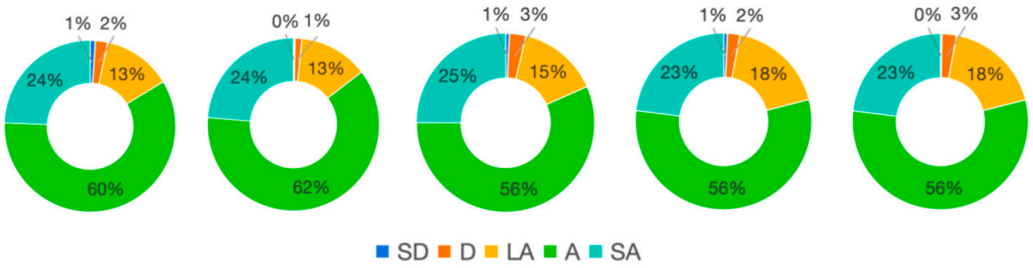
The results of the perspectives on the visual modality in different types of learning applications in Figure 4 showed that a larger proportion of the respondents found the visual modality (e.g., pictures, images, or videos) to be more useful than other modalities in major-based applications because this item scored the highest mean of 4.1 and the standard deviation was 0.73. The item "I find that the visual modality (e.g., pictures, images, videos) is more useful than other modalities in game-based applications" came second in this dimension (M = 4.07, S*D = 0.67). These results indicated that most users believed that the visual mode was more crucial than the other four modalities in major-based, game-based, and function-based applications.

The university students' perspectives in Figure 5 on the scale of the aural modality in different learning applications were mostly agreement and strong agreement. The highest percentage of agreement and highest mean occurred for the item "I find that the aural modality (e.g., audio or listening function) is more useful than other modalities in exam-based applications" (62%, M = 4.07, S*D = 0.67). The item "I find that the aural modality (e.g., audio or listening function) is more useful than other modalities in game-based applications" had the greatest proportion of strong agreement among the five items (25%, M = 4.02, S*D = 0.76). Furthermore, the item "I find that the aural modality (e.g., audio or listening function) is more useful than other modalities in major-based applications" scored second for the mean (M = 4.04, S*D = 0.75). These results indicated that the aural modality was more appreciated in exam-based applications and major-based applications than in the other three types.



1. I find the visual modality (e.g. pictures, images or videos) is more useful than other modalities in major-based applications.
2. I find the visual modality (e.g. pictures, images or videos) is more useful than other modalities in exam-based applications.
3. I find the visual modality (e.g. pictures, images, videos) is more useful than other modalities in game-based applications.
4. I find the visual modality (e.g. pictures, images, videos) is more useful than other modalities in task-based applications.
5. I find the visual modality (e.g. pictures, images, videos) is more useful than other modalities in function-based applications.

Figure 4. Percentage statistics for perspectives on visual modality in different types of learning applications.



1. I find the aural modality (e.g. audio or listening function) is more useful than other modalities in major-based applications.
2. I find the aural modality (e.g. audio or listening function) is more useful than other modalities in exam-based applications.
3. I find the aural modality (e.g. audio or listening function) is more useful than other modalities in game-based applications.
4. I find the aural modality (e.g. audio or listening function) is more useful than other modalities in task-based applications.
5. I find the aural modality (e.g. audio or listening function) is more useful than other modalities in function-based applications.

Figure 5. Percentage statistics for perspectives on aural modality in different types of learning applications.

The findings in Figure 6 that related to the descriptive analysis of students' perspectives on the verbal modality in different learning applications were much more in agreement than disagreement. The highest percentage of college users agreed and strongly agreed that "I find that the verbal modality (e.g., speaking or sound-recording function) is more useful than other modalities in function-based applications" (81%, M = 4.01, S*D = 0.75). The item "I find that the verbal modality (e.g., speaking or sound-recording function) is more useful than other modalities in game-based applications" scored the lowest mean of 3.94, and the standard deviation was 0.81. Based on the results, the university learners reported that the verbal modality was more useful in function-based applications and less appreciated in game-based applications than the other four modalities.

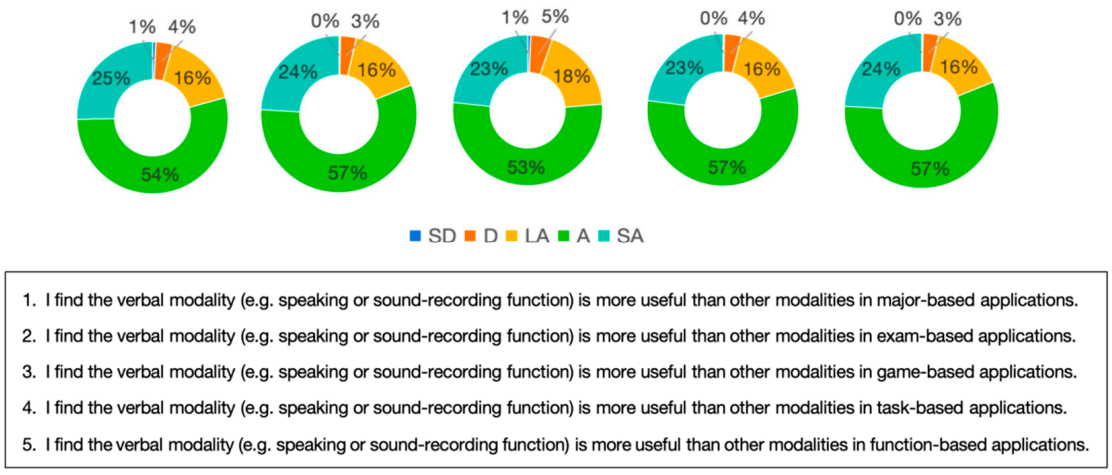


Figure 6. Percentage statistics for perspectives on verbal modality in different types of learning applications.

The findings concerning students' responses in Figure 7 on all items regarding their perspectives on the reading modality in different learning applications indicated that two of the items had the same mean score (4.04) and were related to the reading modality (e.g., text) in major-based applications and exam-based applications. Participants believed that the reading modality was more important in these two applications than in other modes. A total of 3% of the college users strongly disagreed with the statement "I find that the reading modality (e.g., text) is more useful than other modalities in game-based applications" (M = 3.9, S*D = 0.88), which scored first in strong disagreement among the five items under this dimension. According to the findings, the reading modality in major-based applications and exam-based applications was more significant to users, while, in game-based applications, it was less appreciated than in the other four types of applications.

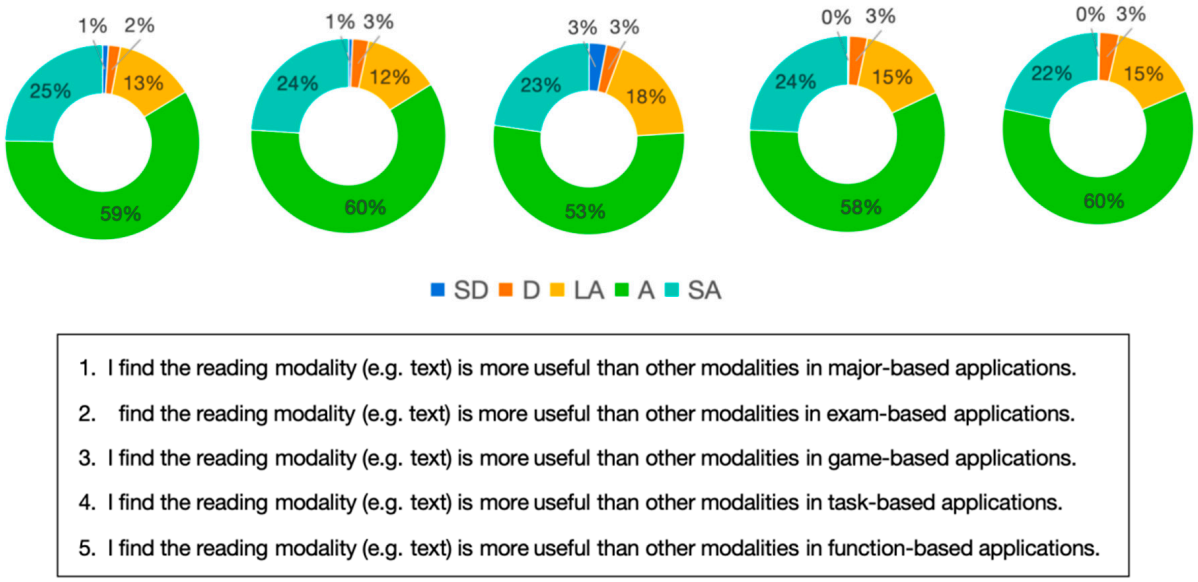


Figure 7. Percentage statistics for perspectives on reading modality in different types of learning applications.

The descriptive analysis of the writing modality scale in Figure 8 illustrated that a large percentage of college learners agreed and strongly agreed with the five items related to the usefulness of this mode. However, around 5% of the respondents disagreed and strongly disagreed here, which was higher than that for the other four modalities. The item "I find

that the writing modality (e.g., word-typing function) is more useful than other modalities in game-based applications" scored highest in disagreement and strong disagreement among the five statements (7%, M = 3.86, S*D = 0.86). Further, the highest mean occurred for the statement "I find that the writing modality (e.g., word-typing function) is more useful than other modalities in exam-based applications" (M = 3.99, S*D = 0.75). Based on these data, the writing modality in exam-based applications was considered more useful than other modalities for college learners, followed by major-based applications and task-based applications, while the writing mode in game-based applications was less useful than the other four types of applications, according to the participants' responses.

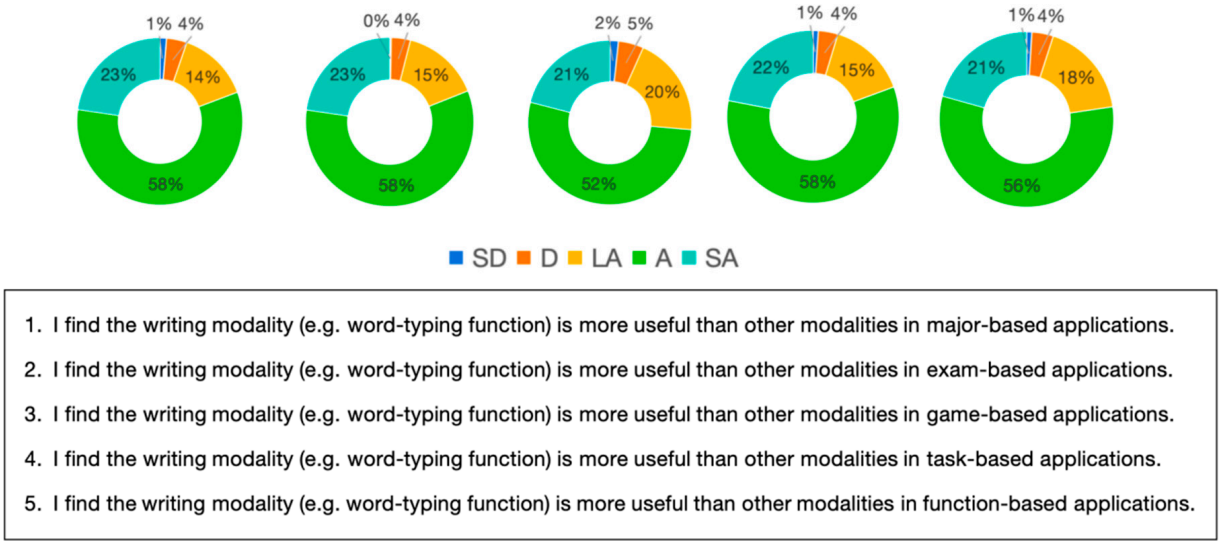


Figure 8. Percentage statistics for perspectives on writing modality in different types of learning applications.

8.2.4. Perceived Satisfaction of Multi-Modalities

The instrument used was adapted from the Post-Study Usability Questionnaire. The PSSUQ is a 16-item survey that measures users' perceived satisfaction with a product or system [44]. Using this scale, students' perceived satisfaction with multi-modalities was reported. The majority of the participants were in agreement and strong agreement with the 20 items in this scale, with slight differences. Over 190 respondents agreed with 35% of the statements. The item "The information shown in multi-modalities was effective in helping me complete tasks and scenarios" scored highest in agreement (66%, number of people: 198) and had the highest mean score of 4.07 (S*D = 0.64), while the item "I find that all the existing modalities in the current game-based applications can satisfy most of my needs to assist my learning performance" had the lowest agreement (56.7%, number of people: 170) and the lowest mean score of 3.87 (S*D = 0.79) of the 20 statements. Thirty percent of the statements had a mean score over 4.0, and the total mean score on this scale was 3.98 with a standard deviation score of 0.72. The corresponding results in the table indicated that multi-modalities were effective in helping university learners to complete tasks and scenarios, and all the existing modalities in the current applications could satisfy most of the users' needs to assist their learning performance, except for game-based learning applications. Most learners agreed that multi-modalities in the learning applications can improve their learning performance with well-organized content and a pleasant interface. Therefore, a high satisfaction with multi-modalities was expressed.

8.2.5. The Relationship Between Multi-Modalities and the Dependent Variable

Table 5 shows all the significant associations between the multi-modality sub-scale and the learners' assisted learning performance through more capable others with the use of mobile technology (applications). The independent variables found to be significantly related to the dependent variable are listed in the title row of the correlation matrix. The rest of the matrix indicates the strength of the relationships among the other variables.

Table 5. Summary of correlation matrix of multi-modalities and learners' assisted learning performance through more capable others with the use of mobile applications.

	Assisted Learning Performance from More Capable Others with the Use of Mobile Technology	**Multi-Modalities**	**Perspectives on Multi-Modalities**	**Visual Modality**	**Aural Modality**	**Verbal Modality**	**Reading Modality**	**Writing Modality**	**Perceived Satisfaction of Multi-modalities**
Assisted learning performance from more capable others with the use of mobile technology	**1**	0.766 **	0.712 **	0.706 **	0.710 **	0.707 **	0.589 **	0.576 **	0.610 **
Multi-modalities	0.766 **	1	0.875 **	0.831 **	0.869 **	0.833 **	0.784 **	0.791 **	0.860 **
Perspectives on multi-modalities	0.712 **	0.875 **	1	0.779 **	0.765 **	0.739 **	0.669 **	0.647 **	0.631 **
Visual modality	0.709 **	0.831 **	0.779 **	1	0.768 **	0.752 **	0.531 **	0.565 **	0.607 **
Aural modality	0.710 **	0.869 **	0.765 **	0.768 **	1	0.820 **	0.642 **	0.633 **	0.654 **
Verbal modality	0.707 **	0.833 **	0.739 **	0.752 **	0.752 **	1	0.619 **	0.614 **	0.607 **
Reading modality	0.589 **	0.784 **	0.669 **	0.531 **	0.642 **	0.642 **	1	0.777 **	0.602 **
Writing modality	0.576 **	0.791 **	0.647 **	0.565 **	0.633 **	0.614 **	0.777 **	1	0.617 **
Perceived satisfaction of multi-modalities	0.610 **	0.860 **	0.631 **	0.607 **	0.654 **	0.607 **	0.602 **	0.617 **	1

Note. **. Correlation is significant at the 0.05 level (2-tailed).

As shown in the table, Pearson's r was used to examine the relationship between multi-modalities and learners' assisted learning performance through more capable others with the use of mobile technology (applications). The findings indicated a positive and strong linear relationship between multi-modalities and learners' assisted learning performance through more capable others with the use of mobile applications [r = 0.766, $p < 0.05$]. Hence, learners' assisted learning performance through more capable others with the use of mobile applications may increase when their perspectives on multi-modalities increase, indicating that there will be an increase in a positive direction in learners' assisted learning performance through more capable others with the use of mobile applications if there is an increased attitude toward multi-modalities in using learning applications. To establish significant relationships between multi-modalities and learners' assisted learning performance through more capable others with the use of mobile applications, respectively, and the link between each of the sub-scales, the Pearson product–moment correlation coefficient was also used.

All seven sub-scales were found to have a significant and strong relationship with the dependent variable: perspectives on assisted learning performance through multi-

modalities (r = 0.712, $p < 0.05$), visual modality (r = 0.709, $p < 0.05$), aural modality (r = 0.710, $p < 0.05$), verbal modality (r = 0.707, $p < 0.05$), reading modality (r = 0.589, $p < 0.05$), writing modality (r = 0.576, $p < 0.05$), and perceived satisfaction with multi-modalities (r = 0.610, $p < 0.05$). The association between these variables and the dependent variable was regarded as strong and significant. The results of the correlation coefficient analysis indicated that learners' assisted learning performance through more capable others with the use of mobile applications was significantly correlated with the seven dimensions under a multi-modality scale. These results indicated that, when college learners have affirmative attitudes toward multi-modalities, their assisted learning performance through more capable others with the use of mobile applications will have affirmative results. According to the results, a positive relationship in the perspectives on multi-modalities, visual modality, aural modality, verbal modality, reading modality, writing modality, and the perceived satisfaction with multi-modalities indicated the positive direction of assisted learning performance through more capable others with the use of mobile applications by these college learners.

9. Discussion

9.1. Finding 1: Assisted Learning Performance Through Multi-Modalities

In this study, research question 1—learners' perspectives on assisted learning performance through multi-modalities in mobile learning applications—was explored both qualitatively and quantitatively. It was based on theory of the Zone of Proximal Development (ZPD) [28], which posits that higher mental functions are built upon lower mental functions and are embedded in the context of the sociocultural milieu; thus, only humans can react indirectly to the world by relying on mediating tools [21]. On the other hand, it was pointed out by [45], for Vygotsky and his followers, the definition of mediation also refers to using "tools". Tools, in this sense, mean anything that is utilized to help solve a problem or achieve a goal. Mobile technology is seen as a mediating tool that can help in the transformation and manipulation of learning activity. Therefore, if appropriate tools based on technology can be used, advanced learning performance can be assured.

The analysis of learners' responses regarding their perspectives on assisted learning performance through multi-modalities in mobile learning applications indicated that students generally enjoyed using various modalities and perceived them as beneficial for enhancing their learning performance, which also agrees with the answers from the interviews. Among the five modalities, the visual modality was the most favored, while the aural modality was considered the most important for learning improvement. In terms of mean scores, the highest was recorded for the perceived importance of the aural modality (M = 4.13, S*D = 0.69), whereas the lowest scores were observed for appreciation the verbal modality (M = 3.99, S*D = 0.79) and the perceived importance of the writing modality (M = 3.99, S*D = 0.75). The overall mean score across all modalities was 4.04 (S*D = 0.74), indicating a generally positive perception of multi-modal learning. These findings suggest that, while students recognize the value of different learning modalities, their preferences and perceived effectiveness vary. The visual modality was the most well received, the aural modality was viewed as essential for learning enhancement, and the verbal and writing modalities were considered less enjoyable and less significant to assisting their learning performance, respectively. This finding also agreed with the second discovery from the interview analysis, which indicated that the visual modality was more appreciated by users in their most-often-used learning apps for many reasons, such as the vivid, attractive, colorful, and straightforward effects.

9.2. Finding 2: The Most Useful Modalities of Mobile Learning Apps

For research question 2, most respondents agreed that they enjoyed using the visual modality (e.g., pictures, images, or videos) in learning applications, and that the aural modality (e.g., audio or listening function) in learning applications was important in enhancing their learning performance. This result agrees with the research of [46,47], who demonstrated that combining text and pictures works best if the pieces of information presented complement each other. This result also conforms with the findings that many students preferred the visual mode in learning applications. The verbal and writing modalities in the learning applications were less enjoyed and less connected with enhancing the learning performance.

Regarding the perspectives on different modes in different types of learning applications (major-based, exam-based, game-based, task-based, and function-based applications), the findings indicated that more respondents found the visual modality (e.g., pictures, images, or videos) more useful than the other modalities in major-based applications; the aural modality (e.g., audio or listening function) was more useful in exam-based applications; the verbal modality (e.g., speaking or sound-recording function) was more useful in function-based applications; the reading modality (e.g., text) was more useful in major-based applications and exam-based applications; and the writing modality (e.g., word-typing function) was more useful in exam-based applications. In these findings, game-based and task-based applications were not mentioned, and exam-based applications were highly preferred among university students, meaning all five modalities in this research were more preferred by users in exam-based apps than in others. Compared with the interview findings, the two results are highly relevant and similar because language learning apps or English exam apps appeared repeatedly in the college students' comments, and they believed that, in these exam-based or function-based apps, all five modalities were important, especially the visual mode and aural mode.

9.3. Finding 3: Perceived Satisfaction with Multi-Modalities

In terms of research question 3, most users believed that multi-modalities were effective in helping them to complete tasks, and all the existing modalities in the current applications satisfied most of the users' needs in assisting their learning performance, except for game-based learning applications. This finding is surprisingly different from many researchers' who aimed to assess the effectiveness of using digital game-based applications to improve learning performance, such as the studies of [48,49]. Their research findings all agreed that the learning motivation of game-based application users improved, and their learning performance was also assisted to some extent. Therefore, the reason behind this phenomenon among Chinese university students leaves a research gap. Under these circumstances, it represents the slow development of game-based learning applications and the under-use or ineffectiveness of the gamification of learning among college students in China. Additionally, most learners agreed that multi-modalities in learning applications can improve their learning performance with well-organized content and a pleasant interface. Overall, a high satisfaction with multi-modalities was expressed.

9.4. Finding 4: The Relationship Between Multi-Modalities and Learners' Assisted Learning Performance Through More Capable Others

The findings for research question 4 revealed a strong positive relationship between the multi-modality sub-scale and learners' assisted learning performance through more capable (knowledgeable) others with the use of mobile technology (applications). There were seven dimensions under the multi-modality sub-scale: perspectives on multi-modalities, perspectives on the visual modality, perspectives on the aural modality, perspectives on the

verbal modality, perspectives on the reading modality, perspectives on the writing modality, and the perceived satisfaction with multi-modalities. Among them, the significant and strong relationship between perspectives on multi-modalities and the dependent variable was the highest ($r = 0.714$, $p < 0.05$), while the perspectives on the writing modality was the lowest ($r = 0.569$, $p < 0.05$). The results showed that positive attitudes or perspectives toward different modalities can significantly influence the effect of learning performance, which is similar to the conclusion of research [50], which showed that, when students hold positive attitudes, they are more likely to feel more motivated and engage more in the learning process, which is an important predictor of successful learning. On the other hand, the participants in this study held the view that, for assisted learning performance with mobile applications, the writing mode is less vital compared with the other modes. This is understandable because the writing mode was barely mentioned in relation to effective learning or learning improvement in the principles of multimedia learning summarized by [47].

10. Conclusions

Based on the results of this study and the extensive discussion, key findings emerged. Most university learners said that their learning performance could be assisted by more capable (knowledgeable) others when they used mobile learning applications; however, they believed that not enough support and guidance was offered. Additionally, the learning apps on the market cannot meet all learners' needs, which calls for the development of new learning applications and more comprehensive functions in the existing learning apps. The results of this study also revealed that the most frequently used learning apps were function-based applications, and a majority of university learners in this study were satisfied with the multi-modalities in the different types of applications, except for game-based apps, to assist their learning performance, which contrasted with the results of several related studies.

From a theoretical perspective, the Zone of Proximal Development theory and multi-modality concept offered valuable insights into the assisted learning performance using mobile learning applications. Zone of Proximal Development (ZPD) theory posits that learning occurs most effectively when a learner is guided by a more knowledgeable other, such as a teacher, peer, or expert, to bridge the gap between what they can accomplish independently and what they can achieve with assistance. The present study aligns with the principles of the Zone of Proximal Development theory by demonstrating that students perceive mobile learning applications as valuable tools for assisted learning, particularly when they provide opportunities for guidance from teachers, peers, or digital mentors. However, the findings also indicate that students believe that existing mobile learning applications do not provide sufficient support, reinforcing the importance of structured scaffolding mechanisms within digital learning environments. Therefore, mobile applications designed for learning should incorporate interactive scaffolding features, such as AI-powered tutoring, real-time expert feedback, and peer collaboration forums, which can enrich prospects for the future Zone of Proximal Development theory.

The multi-modality concept, rooted in cognitive and multimedia learning theories, emphasizes that individuals learn more effectively when exposed to multiple modes of information, such as text, visuals, audio, and interactive elements. Our study's findings reinforce the idea that university students appreciate and benefit from multi-modal learning applications, as they provide a richer and more engaging educational experience. The use of visual (images, diagrams, videos), auditory (audio explanations, podcasts), and textual (reading materials, notes) modes within mobile learning applications contributes to the enhanced comprehension and retention of information. Therefore, integrating multiple

sensory channels to cater to different learning styles, cognitive preferences, and personalized multi-modal learning experiences unites all future branches in the new creative multi-modality model. The findings of the current research contribute to the larger body of knowledge and provided insights into the Zone of Proximal Development theory and multi-modality concept regarding mobile learning technology (applications) use as a whole.

11. Suggestions and Future Directions

Since students believe they do not receive enough guidance when using mobile learning applications, educators should integrate structured support mechanisms such as virtual office hours, discussion forums, and synchronous interactive sessions within mobile learning platforms. Blended learning models that combine mobile applications with teacher-led instruction may also help bridge the support gap. In addition, recommending well-designed applications that align with students' learning objectives and cognitive preferences can also maximize their effectiveness. Educators can implement collaborative learning strategies by integrating mobile applications that allow group discussions, peer feedback, and teamwork.

Given that existing learning applications do not fully meet learners' needs, developers should focus on designing more adaptive and personalized learning systems. Features such as AI-driven content recommendations, customizable interfaces, and multi-modal interaction options (visual, auditory, and textual) can enhance user experience and learning effectiveness. On the other hand, since the students found multi-modal features beneficial in learning applications, developers should ensure that apps support diverse sensory engagement, such as interactive visuals, voice-based explanations, and text-to-speech functionalities. Providing the seamless integration of different modes can enhance accessibility and accommodate various learning styles.

While this study focused on students' perspectives, future research should explore mobile technology use among teachers. Investigating teacher adoption, facilitating conditions, and barriers to mobile technology integration can provide a more comprehensive view of its effectiveness. Observational studies and document analysis can complement self-reported data, offering deeper insights into how mobile learning tools are utilized in real classroom settings.

Secondly, since this research was conducted in universities, further studies should examine mobile learning across different educational levels, including primary and secondary schools. By extending the sample to younger students and teachers in diverse school environments, researchers can better understand how mobile technologies support learning performance across different stages of education. This would also allow for a comparative analysis of mobile learning adoption in various school settings, both within and outside of China.

Additionally, given the varying perceptions of the different modalities in this study, further research can explore why some students prefer certain modalities over others. For instance, experimental or longitudinal studies can assess the effectiveness of different multi-modal approaches in mobile learning applications. Additionally, the role of personal learning preferences and cognitive styles in determining the effectiveness of multi-modal learning strategies can be explored.

On the other hand, some students indicated that game-based learning applications did not fully meet their learning needs compared to other types of learning applications. Future studies can examine the specific design features of game-based learning applications that may contribute to or limit their effectiveness. Researchers can also explore how multi-modal elements within game-based environments can be optimized to enhance learning performance across different subjects and student demographics.

Another potential research direction is investigating why some students feel they receive insufficient support from their teachers in mobile-assisted learning. Qualitative studies involving student and teacher interviews can help identify the gaps in teacher–student interaction within mobile learning environments. Additionally, research can focus on professional development programs that equip teachers with strategies for effectively supporting students in mobile learning settings.

Lastly, to gain a broader perspective on mobile technology adoption, comparative studies can be conducted across different countries. Examining how various educational systems integrate mobile learning and multi-modal strategies will provide insights into best practices and potential cultural or institutional barriers to adoption.

Author Contributions: Conceptualization, R.Y.; methodology, R.Y.; validation, R.Y., H.A.J. and M.K.O.; formal analysis, R.Y.; investigation, R.Y.; resources, R.Y.; data curation, R.Y.; writing—original draft, R.Y.; writing—review & editing, R.Y., H.A.J. and M.K.O.; supervision, H.A.J. and M.K.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Ethics committee: Ethics Committee for Research Involving Human Subjects, Universiti Putra Malaysia. JKEUPM-2023-1443. 12/2/2024.

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

Data Availability Statement: The datasets used and analyzed during the current study are available from the corresponding author on reasonable request. The data are not publicly available due to privacy and ethical reasons.

Acknowledgments: We extend heartfelt appreciation to our university for providing an exceptional academic environment and invaluable resources that were instrumental in the successful completion of this research. Our sincere thanks go to the students who actively participated in and supported various aspects of this research. Their cooperation, insights, and contributions were integral to the comprehensive understanding and analysis presented in this study. Furthermore, we acknowledge the support from other experts we reached out to who played a significant role in making this research possible. Their guidance and resources greatly contributed to the successful execution of this study.

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

References

1. Rakhmatov, D. Mobile technologies in the higher education system. *Ment. Enlight. Sci.-Methodol. J.* **2021**, *2021*, 182–196.
2. Ahmad, T. Student perceptions on using cellphones as learning tools: Implications for mobile technology usage in Caribbean higher education institutions. *PSU Res. Rev.* **2020**, *4*, 25–43. [CrossRef]
3. Ally, M.; Prieto-Blázquez, J. What is the future of mobile learning in education? *Int. J. Educ. Technol. High. Educ.* **2014**, *11*, 142–151. [CrossRef]
4. Subramaniam, G.K.J.; Harun, R.N.S. Adoption of mobile technology in higher education: Students' perceptions of English language learning using smart phones. *Int. J. Asian Soc. Sci.* **2013**, *3*, 2084–2089.
5. Looi, C.K.; Lim, K.F.; Pang, J.; Koh, A.L.H.; Seow, P.; Sun, D.; Boticki, I.; Norris, C.; Soloway, E. Bridging formal and informal learning with the use of mobile technology. In *Future Learning in Primary Schools: A Singapore Perspective*; Chai, S.C., Lim, P.C., Tan, M.C., Eds.; Springer: Singapore, 2016.
6. Pegrum, M.; Oakley, G.; Faulkner, R. Schools going mobile: A study of the adoption of mobile handheld technologies in western Australian independent schools. *Australas. J. Educ. Technol.* **2013**, *29*, 66–81. [CrossRef]
7. Henderson, S.; Yeow, J. IPad in education: A case study of iPad adoption and use in a primary school. In Proceedings of the 45th Hawaii International Conference on System Science (HICSS), Maui, HI, USA, 4–7 January 2012; pp. 78–87.
8. Krull, G.; Duart, J. Research trends in mobile learning in higher education: A systematic review of articles (2011-2015). *Int. Rev. Res. Open Distrib. Learn.* **2017**, 18. [CrossRef]
9. Cheon, J.; Lee, S.; Crooks, S.M.; Song, J. An investigation of mobile learning readiness in higher education based on the theory of planned behavior. *Comput. Educ.* **2012**, *59*, 1054–1064. [CrossRef]
10. Eisenlauer, V. Multimodality in mobile-assisted language learning. *Commun. Comput. Inf. Sci.* **2014**, *479*, 328–338. [CrossRef]

11. Farrah, M.; Abu-Dawood, A.K. Using mobile phone applications in teaching and learning process. *Int. J. Res. Engl. Educ.* **2018**, *3*, 48–68. [CrossRef]
12. Power, R. Design of mobile teaching and learning in higher education: An introduction. In *Handbook of Mobile Teaching and Learning*; Zhang, Y.A., Cristol, D., Eds.; Springer Nature Singapore Pte Ltd.: Berlin/Heidelberg, Germany, 2019; pp. 3–11. [CrossRef]
13. Cross, S.; Sharples, M.; Healing, G.; Ellis, J. Distance learners' use of handheld technologies. *Int. Rev. Res. Open Distance Learn.* **2019**, *20*. [CrossRef]
14. Kumar, B.; Goundar, M.; Chand, S. Usability guideline for mobile learning applications: An update. *Educ. Inf. Technol.* **2019**, *24*, 3537–3553. [CrossRef]
15. Kumar, B.A.; Mohite, P. Usability guideline for mobile learning apps: An empirical study. *Int. J. Mob. Learn. Organ.* **2016**, *10*, 223–237. [CrossRef]
16. Aresta, M.; Pedro, L.; Santos, C. Mobile learning and higher education: A theoretical overview. *J. Mob. Multimed.* **2015**, *11*, 147–156.
17. Gangaiamaran, R.; Pasupathi, M. Review on use of mobile apps for language learning. *Int. J. Appl. Eng. Res.* **2017**, *12*, 11242–11251.
18. Brantes Ferreira, J.; Zanela Klein, A.; Freitas, A.; Schlemmer, E. Mobile learning: Definition, uses and challenges. In *Increasing Student Engagement and Retention Using Mobile Applications: Smartphones, Skype and Texting Technologies*; Emerald Group Publishing Limited: Bingley, UK, 2013; pp. 47–82.
19. Gonzalez, A.; Hwang, G.J. The impact of mobile learning on student engagement and academic performance: A systematic review. *Comput. Educ.* **2021**, *168*, 104194. [CrossRef]
20. Bernacki, M.L.; Greene, J.A.; Crompton, H. Mobile technology, learning, and achievement: Advances in understanding and measuring the role of mobile technology in education. *Contemp. Educ. Psychol.* **2020**, *60*, 101827–101828. [CrossRef]
21. Albadry, H. Using Mobile Technology to Foster Autonomy Among Language Learners. Ph.D. Thesis, Newcastle University, Newcastle upon Tyne, UK, 2018.
22. Chiu, P.; Huang, Y. The development of a decision support system for mobile learning: A case study in Taiwan. *Innov. Educ. Teach. Int.* **2016**, *53*, 532–544. [CrossRef]
23. Lai, Y.; Saab, N.; Admiraal, W. University students' use of mobile technology in self-directed language learning: Using the integrative model of behavior prediction. *Comput. Educ.* **2022**, *179*, 104413. [CrossRef]
24. Yang, Z. A study on self-efficacy and its role in mobile-assisted language learning. *Theory Pract. Lang. Stud.* **2020**, *10*, 439–444. [CrossRef]
25. Ekoç, A. Mobile language learning applications from the perspectives of adult language learners in Turkey. *Shanlax Int. J. Educ.* **2021**, *9*, 259–264. [CrossRef]
26. Kruchinin, S.; Bagrova, E. Quality of mobile apps for language learning. *SHS Web Conf.* **2021**, *93*, 01009. [CrossRef]
27. Noviana, E.; Erlinda, S.; Novianti, R.; Sari, I.K.; Mulyani, E.A.; Zulkifli, N.; Permana, D. Theoretical study to design digital disaster learning resources for prospective elementary school teachers. In *Proceedings of the 4th International Conference on Progressive Education 2022 (ICOPE 2022)*; Advances in Social Science, Education and Humanities Research; Atlantis Press: Dordrecht, The Netherlands, 2023; pp. 647–653. [CrossRef]
28. Vygotsky, L. *Mind and Society: The Development of Higher Psychological Processes*; Harvard University Press: Cambridge, MA, USA, 1978.
29. Tharp, R.G.; Gallimore, R. *Rousing Minds to Life: Teaching, Learning, and Schooling in Social Context*; Cambridge University Press: Cambridge, UK, 1988.
30. Ab Jalil, H.; McFarlane, A.; Ismail, I.A.; Rahman, F. Assisted performance a pragmatic conception of online learning. *Int. J. Instr.* **2008**, *1*, 57–76.
31. Liu, X.Y.; Meng, P.; Xie, Y.X.; Yu, H.Y. An analysis of multi-modal reading class for elementary English language learners under immersive language environment—Take American classes as an example. *Adv. Soc. Sci. Educ. Humanit. Res.* **2022**, *637*, 805–812. [CrossRef]
32. Baltrušaitis, T.; Ahuja, C.; Morency, L.P. Multimodal machine learning: A survey and taxonomy. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *41*, 423–443. [CrossRef]
33. Guo, Z.; Li, X.; Huang, H.; Guo, N.; Li, Q. Deep learning-based image segmentation on multi-modal medical imaging. *IEEE Trans. Radiat. Plasma Med. Sci.* **2019**, *3*, 162–169. [CrossRef]
34. Tariq, A. Approaches to different learning styles in undergraduate medical students of Al-Tibri medical college Karachi. *Ann. PIMS-Shaheed Zulfiqar Ali Bhutto Med. Univ.* **2023**, *19*, 284–288. [CrossRef]
35. Eshankulovna, R.A. Modern technologies and mobile apps in developing speaking skill. *Linguist. Cult. Rev.* **2021**, *5*, 1216–1225. [CrossRef]

36. Viberg, O.; Kukulska-Hulme, A. Fostering learners' self-regulation and collaboration skills and strategies for mobile language learning beyond the classroom. In *Routledge Handbook of Language Teaching and Learning Beyond the Classroom*; Reinders, H., Lai, C., Sundqvist, P., Eds.; Routledge: New York, NY, USA, 2021; pp. 1–15.
37. Dörnyei, Z.; Taguchi, T. *Questionnaires in Second Language Research: Construction, Administration, and Processing*; Routledge: New York, NY, USA, 2010. [CrossRef]
38. Alqahtani, M.; Mohammad, H. Mobile applications' impact on student performance and satisfaction. *Turk. Online J. Educ. Technol.* **2015**, *14*, 102–112.
39. Bernard, H.R. *Research Methods in Cultural Anthropology*; Sage: Newbury Park, CA, USA, 1988.
40. McLellan, L.E.; MacQueen, K.M.; Neidig, J.L. Beyond the qualitative interview: Data preparation and transcription. *Field Methods* **2003**, *15*, 63–84. [CrossRef]
41. Welsh, E. Dealing with data: Using NVivo in the qualitative data analysis process. *Forum Qual. Sozialforschung Forum Qual. Soc. Res.* **2002**, *3*. [CrossRef]
42. Corbin, J.; Strauss, A. *Basics of Qualitative Research Techniques and Procedures for Developing Grounded Theory*, 3rd ed.; Sage Publications: Newbury Park, CA, USA, 2007.
43. Miles, M.B.; Huberman, A.M. *Qualitative Data Analysis: An Expanded Sourcebook*, 2nd ed.; Sage Publications: Thousand Oaks, CA, USA, 1994.
44. Adrian, G. UX Research: Standardized Usability Questionnaire. Choosing the Right Usability Questionnaire. ChaiOne. 2013. Available online: https://arl.human.cornell.edu/linked%20docs/Choosing%20the%20Right%20Usability%20Questionnaire.pdf (accessed on 6 March 2025).
45. Williams, M.; Burden, R. *Psychology for Language Teachers: A Social Constructivist Approach*; Cambridge University Press: Cambridge, UK, 1997.
46. Mayer, R.E. Cognitive theory of multimedia learning. In *The Cambridge Handbook of Multimedia Learning*; Mayer, R.E., Ed.; Cambridge University Press: Cambridge, UK, 2005; pp. 31–48. [CrossRef]
47. Miladiyah, M. The outcome of multi-modality on language teaching and learning. *Adv. Soc. Sci. Educ. Humanit. Res.* **2018**, *228*, 341–347. [CrossRef]
48. Hartt, M.; Hosseini, H.; Mostafapour, M. Game on: Exploring the effectiveness of game-based learning. *Plan. Pract. Res.* **2020**, *35*, 589–604. [CrossRef]
49. Berns, A.; Isla-Montes, J.L.; Palomo-Duarte, M.; Dodero, J.M. Motivation, students' needs, and learning outcomes: A hybrid game-based app for enhanced language learning. *SpringerPlus* **2016**, *5*, 1305. [CrossRef]
50. Allagui, B. Factors underlying students' attitudes towards multimodal collaborative writing. *Heliyon* **2022**, *9*, e12725. [CrossRef]