Identifying key factors of dynamic ADDIE model for instructional virtual reality design: An exploratory study

Ruo-Yu Li, Jerry Chih-Yuan Sun

Institute of Education, National Yang Ming Chiao Tung University, Hsinchu City, Taiwan

Address for Correspondence: Jerry Chih-Yuan Sun, Institute of Education, National Yang Ming Chiao Tung University, No. 1001, Daxue Rd., East Dist., Hsinchu City 300093, Taiwan. Email: jerrysun@nycu.edu.tw.

Identifying key factors of dynamic ADDIE model for instructional virtual reality design: An exploratory study

Due to the features of immersive learning and without the restrictions of time and place, virtual reality (VR) is commonly used in learning areas. For investigating functional design in the VR learning environment, previous literature proposed a series of theoretical models and strategies without considering the key factors for their development. Therefore, this study explored the key factors based on the dynamic ADDIE model for the design of VR instructional systems. To augment the assessment integrity, this study adopted the fuzzy Delphi method (FDM). Open and closed questions were also integrated into the online FDM-based questionnaire to assess the opinions of expert groups, with 9 professors and 9 school teachers. The results showed that the final selection encompassed 5 criteria, 2 sub-criteria, and 23 key factors. In the overall priority order, the "interactive feedback design" was ranked as the most influential factor, while the "learner opinion survey" was the least. Most importantly, since the experts considered "formative evaluation" to be the highest significant sub-criterion, the reliability of the dynamic ADDIE model for instructional VR proposed in this study was confirmed. In addition, the divergent opinion from expert groups could provide references for future developers with different backgrounds.

Keywords: dynamic ADDIE model; FDM, fuzzy Delphi method; VR design; VR learning

Introduction

In recent years, virtual reality (VR) technology has been widely applied in learning areas. This phenomenon can be explained due to its advantages of not only being unlimited by time and place, but also its intriguing user experience and immersive learning environment (Makransky et al., 2020;

Parong & Mayer, 2021; Wang et al., 2018). Furthermore, previous studies showed that VR could increase students' learning interest, learning motivation, and learning outcome (Fowler, 2015; Hu-Au & Lee, 2017; Petersen et al., 2020). However, studies based on the design research method mostly summarized design factors through literature reviews and personal experiences during development (Bogusevschi et al., 2020; Buche et al., 2010; Fiard et al., 2014), which involved reliability and validity concerns.

For designing VR instructional systems, previous studies adopted design-based research or incorporated generative learning strategies into the learning process (Cochrane et al., 2017; Makransky et al., 2021). According to Warren et al. (2014), instructional design factors such as students' and teachers' beliefs in knowledge are strongly impacted by technology. As a result, the ADDIE model comes into play and has been examined by prior studies (Gusmida & Islami, 2017; Trust & Pektas, 2018; Wiphasith et al., 2016), and it is the most commonly used model in the instructional design for virtual environments (Göksu et al., 2017; Soto, 2013). There were several revisions of the original ADDIE model for more dynamic features (Allen, 2006; Dick et al., 2009; Smith & Ragan, 2004); nevertheless, a limited body of literature explored the feasibility of the dynamic ADDIE model utilized in the education field.

To augment reliability and validity, the fuzzy Delphi method (FDM) enables scholars to collect the opinions of experts from diverse fields in a more efficient way (Murray et al., 1985). Furthermore, by employing the corresponding variables from expert opinions, the evaluation process of the factors is more objective (Noorderhaben, 1995). Thus, this study used FDM to determine the feasibility of potential factors that contribute to the dynamic ADDIE model. Among the educational literature, Li and Wang (2020) conducted FDM research for the location-based game design but could not fully account for the results based only on the quantitative data. Therefore, to improve the deficiency of explanation, this study applied open- and closed-ended questions to obtain quantitative and qualitative data.

In conclusion, the primary research questions in this study are as follows:

(1) What are the key factors for the design of an instructional VR environment?

(2) What is the design model based on the dynamic ADDIE for instructional VR?

(3) What are the reasons for the composition of the priority of key factors?

Literature review

This study reviewed the literature associated with the ADDIE model and instructional VR systems and ultimately identified five criteria and 25 initial factors, as presented in Table 1 at the end of this section.

Analysis

The analysis criterion refers to the assessment of the learning experience and requirements using VR, then the setting of the instructional goal. Stavroulia et al. (2019) and Yu et al. (2021) both adopted a focus group interview to understand the required aspects, anticipation, and experience of their participants. Furthermore, Sholihin et al. (2020) performed an analysis by observing in the classroom. Another approach commonly used was a questionnaire survey to obtain development needs from users (Akman & Çakır, 2020; Stavroulia et al., 2019). Accordingly, three factors were defined in this criterion, namely “focus group interview,” “learner status observation,” and “learner opinion survey.”

Design

The design criterion refers to the appropriate VR learning content that is well designed. Given that learning content frequently involves professional knowledge, the search for the expert opinion was extensively used during the design phase (Akman & Çakır, 2020; Asad et al., 2021; Sholihin et al., 2020; Stavroulia et al., 2019; Yu et al., 2021). For the other methods, Alrehaili and Osman (2022) applied the multimedia principles proposed by Clark and Mayer (2016) to design their VR game; moreover, Sholihin et al. (2020) arranged a storyboard for their content design of learning media. Referring to related materials for VR instructional design was indeed another useful way that was warranted (Alrehaili & Osman, 2022; Stavroulia et al., 2019). Thus, the “expert opinion survey,” “multimedia principles,” “storyboard,” and “data collection” are defined as four factors in this criterion.

Development

The development criterion refers to the construction of a VR instructional system. On the one hand, several decisions are made during this phase. First, the selection of a type of VR, such as immersive VR, is made, depending on limited resources and funds (Sholihin et al., 2020). Second, a development tool such as Unity is decided on based on its usability and cross-platform support (Astuti et al., 2020). Third, a VR display, such as HTC Vive, is chosen according to immersion level (Stavroulia et al., 2019). On the other hand, 3D modeling, illustration and video arrangement, and interactive feedback design play essential roles in the development (Akman & Çakır, 2020; Yu et al., 2021). In light of this, researchers defined six factors in this criterion, namely "content and resource assessment," "VR development tool assessment," "VR display assessment," "3D modeling," "media integration," and "interactive feedback design."

Implementation

The implementation criterion refers to the preparation and execution of a VR environment where learners are formally engaged in learning. Before the course was delivered, Yu et al. (2021) applied a prior knowledge test and assisted the learners by providing VR instruction. When the course started, Akman and Çakır (2020) provided voice guidance during their VR game to aid in task understanding. Similarly, Asad et al. (2021) evaluated the learning results at each stage to identify their problems and possible solutions. Furthermore, Stavroulia et al. (2019) used fitness wristbands and an electroencephalogram (EEG) to record participants' heart rates and brain signals. Hence, five factors in this criterion were defined: "prior-knowledge test," "pre-class instruction," "in-class instruction," "formative assessment," and "physiological data measurement."

Evaluation

The evaluation criterion encompasses two sub-criteria. One is formative evaluation, which refers to analysis and testing after the first four phases of the ADDIE model. Asad et al. (2014) took the evaluation after the analysis phase, for instance, which could discover and solve problems of the ongoing process. Similarly, the evaluation also served as a way to detect potential system problems at the end of VR development (Astuti et al., 2020). Therefore, four factors are literally defined as "analysis phase evaluation," "design phase evaluation," "development phase evaluation," and "implementation phase evaluation" in this criterion.

Another sub-criterion is the summative evaluation, which refers to the evaluation of learning outcomes and feedback from learners after VR learning. Indeed, posttests to estimate learning achievement and cognitive status were widely adopted in the research field (Chang et al., 2019; Sholihin et al., 2020; Stavroulia et al., 2019); in addition, it is significant to know learners' satisfaction and suggestions for future development (Yu et al., 2021). Thus, researchers have defined the "learning achievement test," "learner satisfaction survey," and "psychological assessment scales" as the three factors that make up summative evaluation. In general, the evaluation phase determined by this study aligns with the dynamic ADDIE model (Allen, 2006).

Table 1. Initial factors for the design of the VR instructional system.

Criterion		Factor		Definition	Reference
C1	Analysis	f1	Focus group interview	Interview learners about their experiences of VR and their learning goals.	Stavroulia et al., 2019; Yu et al., 2021
		f2	Learner status observation	Observe how learners learn with VR tools.	Sholihin et al., 2020
		f3	Learner opinion survey	Conduct a survey to understand what learners attempt to learn and their goals.	Akman & Çakır, 2020; Stavroulia et al., 2019

C2	Design	f4	Expert opinion survey	Collect opinions with reference to expert interviews or surveys on VR learning content.	Akman & Çakır, 2020; Asad et al., 2021; Sholihin et al., 2020; Stavroulia et al., 2019; Yu et al., 2021
		f5	Multimedia principles	Refer to Mayer's multimedia principles, e.g., streamline textual narratives with the coherence principle.	Alrehaili & Osman, 2022
		f6	Storyboard	Decompose the video media embedded in VR into a series of sketches to present the order.	Sholihin et al., 2020
		f7	Data collection	Search for educational materials or literature related to the construction of the VR system.	Alrehaili & Osman, 2022; Stavroulia et al., 2019
C3	Development	f8	Content & Resource assessment	Evaluate learning content, limited resources, and funding to develop a desktop, projection or immersive VR instructional system.	Sholihin et al., 2020
		f9	VR development tool assessment	Assess the usability and cross-platform support of VR development tools, then choose a suitable one, e.g., Unity.	Asad et al., 2021; Astuti et al., 2020; Stavroulia et al., 2019; Yu et al., 2021
		f10	VR display assessment	Select VR displays according to the immersion level, e.g., HTC Vive.	Stavroulia et al., 2019; Yu et al., 2021
		f11	3D modeling	Build 3D objects in a VR instructional system to create a fully realistic environment.	Asad et al., 2021; Astuti et al., 2020; Sholihin et al., 2020; Stavroulia et al., 2019; Yu et al., 2021
		f12	Media integration	Add videos and pictures to a VR instructional system.	Sholihin et al., 2020; Yu et al., 2021
		f13	Interactive feedback design	Design the corresponding feedback for body movements, focus, etc. of the learners.	Akman & Çakır, 2020; Astuti et al., 2020; Yu et al., 2021
C4	Implementation	f14	Prior-knowledge test	Examine learners' prior knowledge of VR learning content to adjust assistance levels.	Yu et al., 2021
		f15	Pre-class instructions	Assist learners to use and test the VR smoothly and understand the flow before the lesson.	Yu et al., 2021
		f16	In-class instructions	Provide immediate guidance and feedback to learners during the learning process.	Akman & Çakır, 2020; Yu et al., 2021
		f17	Formative assessment	Incorporate an assessment into the VR learning process and provide instructions after the test.	Asad et al., 2021
		f18	Physiological data measurement	Use wearable devices to measure physiological data to investigate learners' emotions and concentration levels.	Stavroulia et al., 2019
C5	Evaluation: Formative	f19	Analysis phase evaluation	Collect the results of the survey or observation to analyze the problems and propose solutions.	Asad et al., 2014
		f20	Design phase evaluation	Evaluate whether the VR design achieves the teaching objectives and meets the needs of the learners.	Asad et al., 2014
		f21	Development phase evaluation	Test the VR system at the end of the development phase before implementation.	Asad et al., 2014; Astuti et al., 2020
		f22	Implementation phase evaluation	Evaluate the pros and cons of the use of VR by learners in the implementation phase and try to improve them.	Asad et al., 2014; Astuti et al., 2020
	Summative	f23	Learning achievement test	Perform achievement tests before and after VR learning to assess learning outcomes.	Chang et al., 2022; Stavroulia et al., 2019; Yu et al., 2021
		f24	Learner satisfaction survey	Collect feedback from learners on their satisfaction, insights, and suggestions.	Chang et al., 2022; Sholihin et al., 2020; Yu et al., 2021
		f25	Psychological assessment scales	Measure the cognitive status, e.g., learning motivation, with scales.	Chang et al., 2022; Sholihin et al., 2020

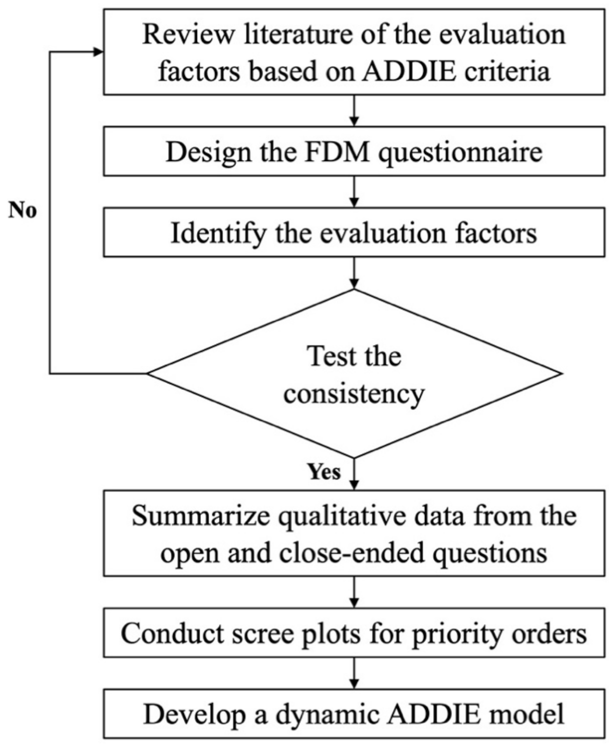
Method

Proposed procedure

This study proposed a combined method using FDM, as demonstrated in Figure 1. To streamline the qualitative survey, the researchers incorporated open- and closed-ended questions into the FDM questionnaire. The first procedure was a review of the literature that aimed to explore the initial factors according to five criteria based on the dynamic ADDIE model to develop questionnaire items. After receiving responses from the online Qualtrics questionnaire, this study used two triangular fuzzy numbers and a gray zone test proposed by Wei and Chang (2008) to calculate consensus values for consistency evaluation.

The researchers subsequently summarized the answers to the open- and closed-ended questions to explore the reasons for the inconsistent factors. Until all values were consistent, screening the initial factors through scree plots was crucial to determine the key factors and priority orders. The researchers then reviewed the qualitative data again for explanations of the FDM results and compared the congruent and incongruent opinions of the two groups of experts. Finally, the key factors under five criteria were chosen to establish the dynamic ADDIE model as a reference for instructional VR design.

Figure 1. The procedure of the proposed FDM method.



Participants

This study invited 18 experts to participate in the investigation by purposive sampling. Although the sample size of Delphi studies has been inconclusive, the literature suggested that if the group of experts was homogeneous, a small sample size of about 8 to 15 may provide sufficient results (Skulmoski et al., 2007; Trevelyan & Robinson, 2015). Consequently, 9 professors and 9 teachers from elementary to high school were recruited. All experts had an average of 3 years of experience conducting research or teaching in related areas, such as VR learning, VR application design, and VR pedagogy employed in VR environments.

FDM questionnaire design

The FDM questionnaire of this study referred to Wei and Chang (2008). The researchers applied Qualtrics to design, distribute, and collect our online questionnaire. In the first part of the questionnaire, the participating experts were asked to give a set of important values for each factor on a scale of 1 to 10. This set of values encompassed three categories: the first was the minimum, mean conservative value, shown as "C^i," indicating the most conservative assessment of the factor importance level given by an expert; the second was the single value, expressed as "a^i," demonstrating a single quantitative value that was equivalent to the ideal importance level; the third was the maximum, that is, the optimistic value, with "O^i," representing the most optimistic evaluation of the factor importance level. Furthermore, the participating experts were also required to check their values following the rule: $C^i < a^i < O^i$.

In the second part, this study organized seven open- and closed-ended questions to collect experts' opinions. Specifically, to identify the reasons for the factor judgment, a multiple-choice question was designed to ask about the conditions that experts consider indispensable to the instructional VR design. Then, according to the priority order given by the experts, they were asked to briefly explain the reasons for the selection of the top three and least three important factors. Lastly, participating experts were asked to determine whether the importance of factors would vary depending on the subject or the topic delivered; furthermore, they could raise possible reasons other than the subject matter.

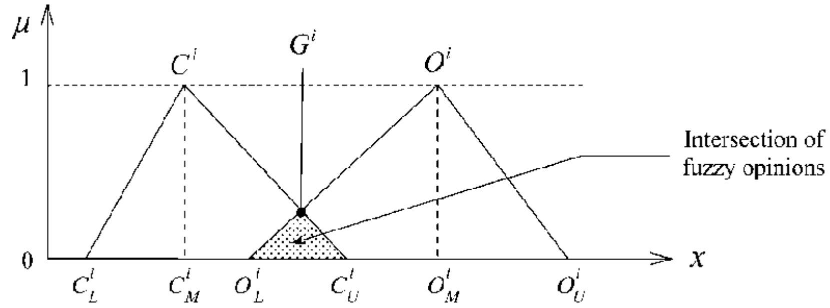
FDM analysis method

Through the FDM data analysis, the extent of the semantic meaning of each expert could be measured. The following paragraphs demonstrate the four analysis steps in this study.

Step 1 was to exclude extreme values and calculate the importance values of C_L^i, C_M^i, C_U^i, O_L^i, O_M^i, and O_U^i. After collecting all data from the questionnaire, extreme values outside the double standard deviation were excluded. Among the remaining values, the researchers calculated the minimum values C_L^i, the geometric means C_M^i, and the maximum values C_U^i in the conservative values. The researchers also reckoned the minimum values O_L^i, the geometric means O_M^i, and the maximum values O_U^i in the optimistic values.

Step 2 was to calculate two triangular fuzzy numbers. Each evaluation factor has two sets of triangular fuzzy numbers, namely (C_L^i, C_M^i, C_U^i) and (O_L^i, O_M^i, O_U^i) from the conservative value C^i and optimistic value O^i respectively. The positions of six values and the gray zone, the intersection of fuzzy opinions in two triangular fuzzy numbers, are illustrated in Figure 2 (Wei & Chang, 2008). Furthermore, the axis μ is the subordinate degree; axis χ is the cognitive value; and G^i is the consensus value.

Figure 2. Two triangular fuzzy numbers. (Wei & Chang, 2008)



Step 3 was to examine the consistency of the consensus values. In this procedure, researchers took the means of a consistency test by estimating the overlap of double triangles, namely the range of geometric means constructed by conservative values and optimistic values. As long as the consensus value could be reckoned with, the opinion of the participating experts showed consistency. To be specific, there were three situations. In the first case, if $C_U^i \leq O_L^i$, the consensus value G^i was equivalent to the arithmetic mean of C_M^i and O_M^i. The equation was displayed as:

$$G^i = \frac{(C_M^i + O_M^i)}{2} \tag{1}$$

In the second case, if $C_U^i > O_L^i$, the fuzzy gray zone was defined as Z^i, wherein $Z^i = C_U^i - O_L^i$. The range of the geometric mean between conservative and optimistic values was defined as M^i, wherein $M^i = O_M^i - C_M^i$. Provided that the requirement of $Z^i < M^i$ was satisfied, the difference between extreme values and other values was not distinct. In this case, the way to find the consensus value G^i was to calculate two triangular fuzzy numbers through the intersection. The maximum of the subordinate degree "μ" was estimated after obtaining a set of fuzzy numbers. Two equations involved in the aforementioned statement are demonstrated as:

$$F^i(\chi_j) = \left\{\int_{\chi} \{\min[C^i(\chi_j),\ O^i(\chi_j)]\}\, dx\right\} \tag{2}$$

$$G^i = \chi_j / \max \mu_{\underset{F}{i}}(\chi_j) \tag{3}$$

In the third case, if $C_U^i > O_L^i$, the participating experts did not reach a consensus. When this situation occurred, the consensus values of the evaluation factors could not be counted. Therefore, more references were needed from participating experts to clarify the vagueness. Most importantly, the questionnaire investigation was implemented continuously until all consensus values were acquired.

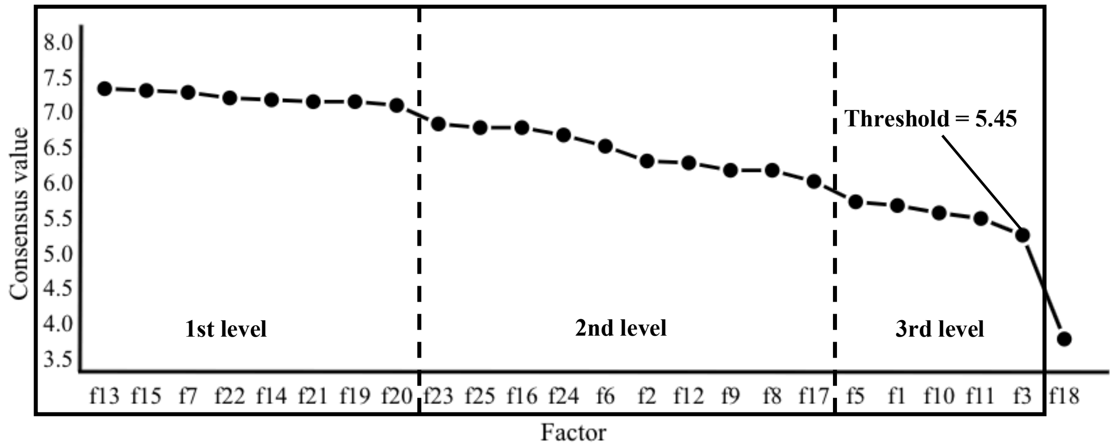
Step 4 was to identify the evaluation factors with scree plots. Researchers ultimately demonstrated the consensus values G^i of all factors in descending order by the linear graph. The threshold of expert consensus was identified by observing the steep slope. Only the consensus values of the factors above or equal to the threshold could be determined as key factors in this study.

Results

Statistics of FDM and scree plots

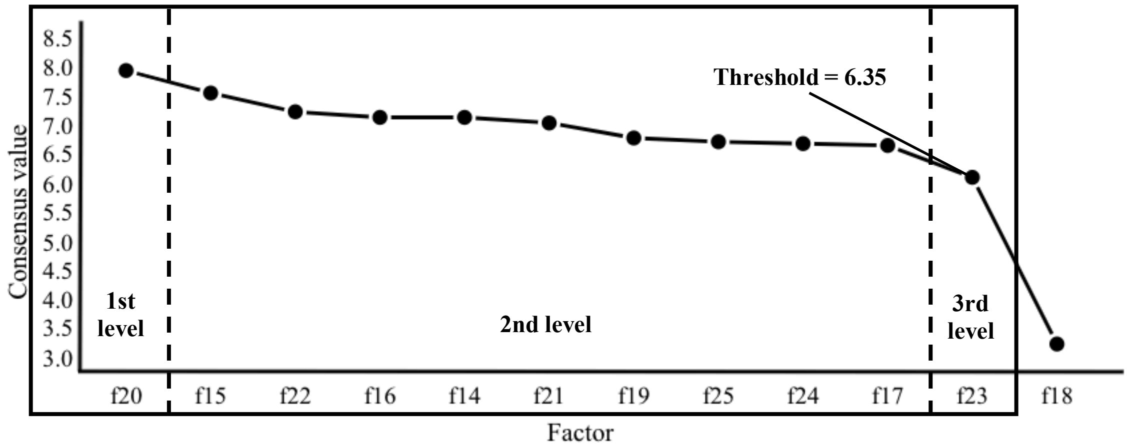
After the second round investigation, all consensus values G^i were obtained excluding f4 "expert opinion survey." As shown in Figure 3, the highest slope in the scree plot was between f3 "learner opinion survey" and f18 "physiological data measurement." Thus, G^i of f3 was set as the threshold value leading to the exclusion of f18, and the remaining 23 items were determined as the key factors of this study. In addition, the key factors were divided into three levels of importance with the second and third steep slopes.

Figure 3. Scree plot of the overall factors.



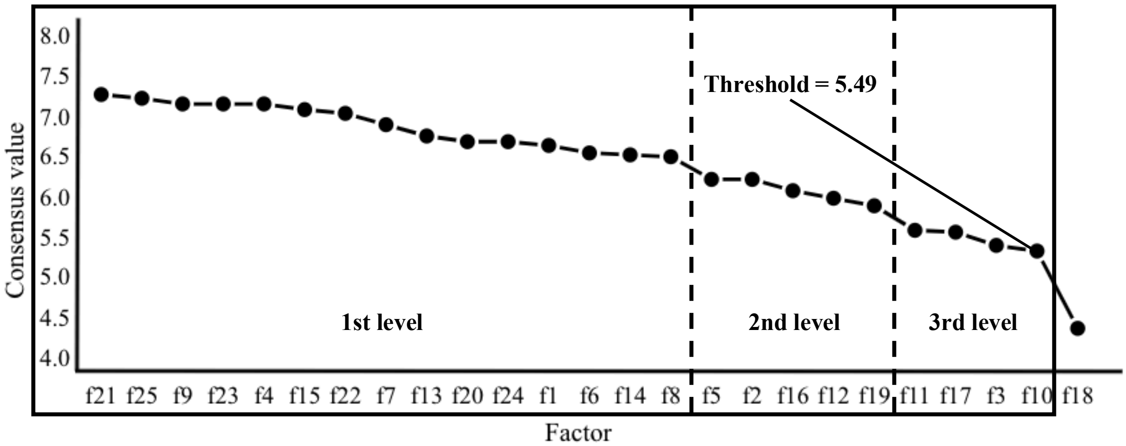
Concerning the teacher group, this study only analyzed factors within the implementation and evaluation criteria. Similarly, the scree plot, as demonstrated in Figure 4, also removed f18.

Figure 4. Scree plot of the teacher group factors.



For the professor group, all factors reached unanimity including f4, and f18 was also excluded, as illustrated in Figure 5.

Figure 5. Scree plot of the professor group factors.



Priority order of the key factors

As shown in Table 2, among the 23 key factors, the participating experts considered that the top three important factors were f13 "interactive feedback design," f15 "pre-class instructions," and f7 "data collection;" the three least important factors were f3 "learner opinion survey," f11 "3D modeling," and f10 "VR display assessment." Furthermore, all the factors that made up the "formative evaluation" were observed at the first level; the other three factors of the "summative evaluation" were less important than the preceding ones. In addition, most of the factors in the development and analysis criteria were ranked lower in priority, and especially the factors in the analysis criterion were almost all classified at the third level.

Table 2. Overall priority order of key factors.

Importance level	Criterion	Factor		G^i	Threshold ($G^i \geq 5.45$)	Overall order
1st level	Development	f13	Interactive feedback design	7.53	7.53 > 5.45	1
	Implementation	f15	Pre-class instruction	7.51	7.51 > 5.45	2
	Design	f7	Data collection	7.47	7.47 > 5.45	3
	Evaluation	f22	Implementation phase evaluation	7.39	7.39 > 5.45	4
	Implementation	f14	Prior-knowledge test	7.36	7.36 > 5.45	5
	Evaluation	f21	Development phase evaluation	7.35	7.35 > 5.45	6
		f19	Analysis phase evaluation	7.34	7.34 > 5.45	7
		f20	Design phase evaluation	7.30	7.30 > 5.45	8
2nd level		f23	Learning achievement test	7.02	7.02 > 5.45	9
		f25	Psychological assessment scales	6.97	6.97 > 5.45	10

	Implementation	f16	In-class instruction	6.97	6.97 > 5.45	11
	Evaluation	f24	Learner satisfaction survey	6.88	6.88 > 5.45	12
	Design	f6	Storyboard	6.72	6.72 > 5.45	13
	Analysis	f2	Learner status observation	6.51	6.51 > 5.45	14
	Development	f12	Media integration	6.47	6.47 > 5.45	15
		f9	VR development tool assessment	6.37	6.37 > 5.45	16
		f8	Content & Resource assessment	6.36	6.36 > 5.45	17
	Implementation	f17	Formative assessment	6.22	6.22 > 5.45	18
3rd level	Design	f5	Multimedia principles	5.92	5.92 > 5.45	19
	Analysis	f1	Focus group interview	5.88	5.88 > 5.45	20
	Development	f10	VR display assessment	5.76	5.76 > 5.45	21
		f11	3D modeling	5.68	5.68 > 5.45	22
	Analysis	f3	Learner opinion survey	5.45	5.45 = 5.45	23

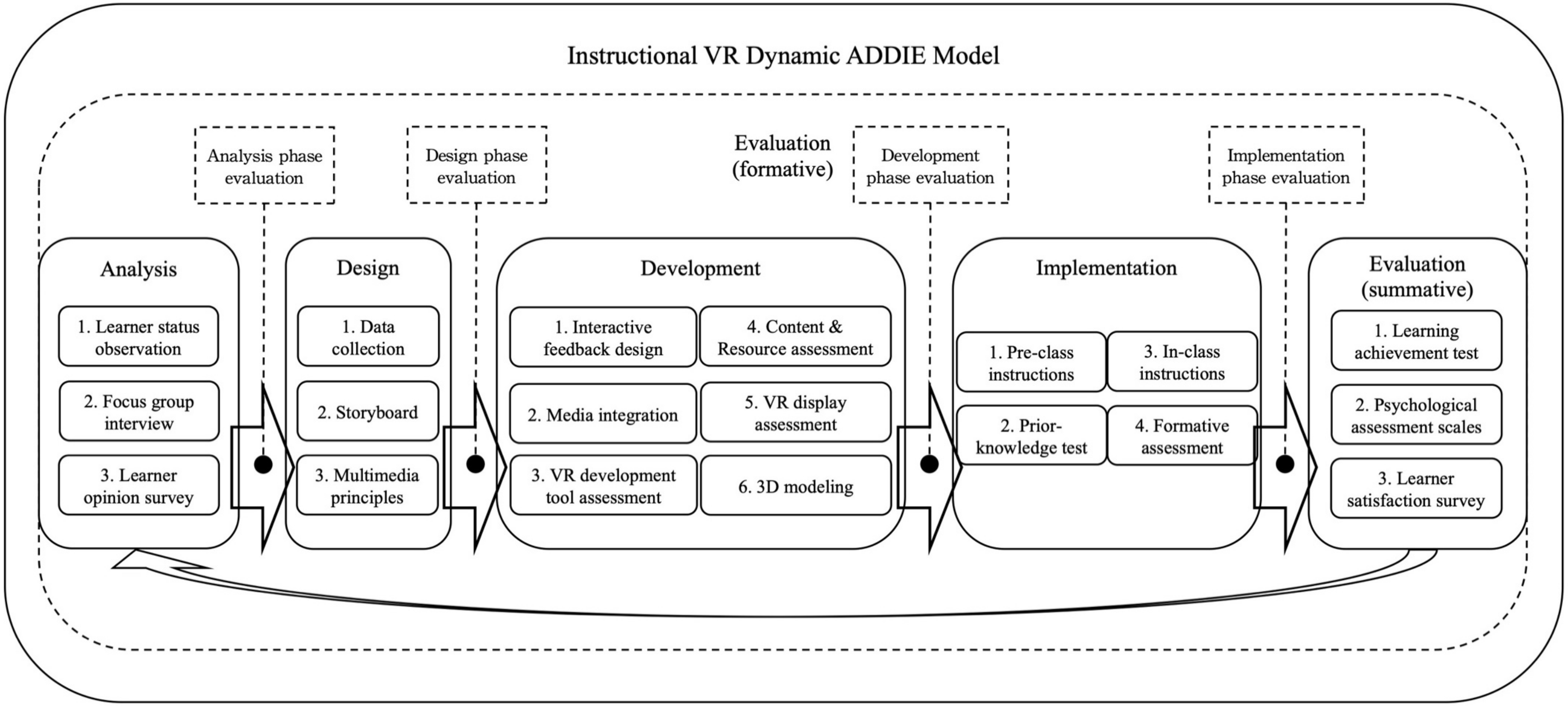
Qualitative data of experts' opinions

In terms of the conditions of the instructional VR design, the "facilitate student learning motivation" was the most popular, selected by 14 experts. Furthermore, 78% of the experts considered that teaching subjects or topics would change the importance of factors. The factors most influenced were f12 "media integration design" and f8 "content and resource assessment." Generally, only f1 "focus group interview" was not considered affected. In addition, five experts provided other reasons, such as: "differences in instructional design, VR usage, and function due to domain distinctions;" "teachers' own knowledge, expertise, and information capabilities;" "the selection of materials that could impact outcomes;" "differential media preferences of the age groups;" and "the length of time and opportunity to learn."

Dynamic ADDIE model for instructional VR design

Based on the key factors of this study, the researchers constructed a dynamic ADDIE model for VR instructional system design, as presented in Figure 6.

Figure 6. Instructional VR dynamic ADDIE model.



Discussion

Reasons for the priority order of the key factors

For the most influential key factor, namely "interactive feedback design," experts in both groups believed that it could enhance student learning interest, reflect learning status, and improve evaluation, which were also selected as the most important conditions for the factors. The following qualitative data from the experts reinforced this result:

> "It could be beneficial for teachers since we were often unable to grasp the concerns of students when learning in a VR environment." (Teacher group, Expert E2)
>
> "Interaction was necessary for evaluation." (Professor group, Expert E13)

On the contrary, the "learner opinion survey" was the least important key factor ranked by both groups. Experts commonly regarded it as a useless approach during the analysis phase. The following qualitative data supported this result:

> "Elementary school students usually could not comprehend the questionnaire items and had little opinion, resulting in low reliability of the survey results." (Teacher group, Expert E6)
>
> "It was not helpful for the practical VR instructional development." (Professor group, Expert E11)

The most critical sub-criterion: formative evaluation

The most important criterion was defined as the percentage of factors in the criterion that was in the first level of the overall priority order. Consequently, "formative evaluation" was considered the most essential sub-criterion since all factors were classified at the first level. This result was verified by the following qualitative data:

> "Design phase evaluation was the basis for the information on systems and materials." (Professor group, Expert E15)
> "The need to assess student learning during the implementation phase was greater at elementary school." (Teacher group, Expert E5)
> "System evaluation in the development phase was able to debug and identify usability problems; moreover, it could review the suitability of the learning content arrangement." (Professor group, Expert E11)

Most importantly, this result provided strong evidence that, instead of the original waterfall ADDIE model, the dynamic features were indispensable. It also aligned with the views from previous studies (Allen, 2006; Asad et al., 2014; Dick et al., 2009; Smith & Ragan, 2004).

Different preferences of the expert groups

The researchers observed some distinctive preferences of the two groups of experts. Comparing factors that differed by more than 10 places between the professor group and the overall priority order, the professor group generally rated the "VR development tool assessment" more important, while the teacher group considered the "analysis phase evaluation" to be more significant. Additionally, "focus group interview" was ranked first in importance by the professor group, but third by the overall experts. These results were corroborated by the following qualitative data:

> "It was vital to know that the VR learning activities required a smartphone, namely mobile VR, or a high-end head-mounted display." (Professor group, Expert E12)
> "Most of the students had not used VR before, which failed to obtain constructive suggestions." (Teacher group, Expert E8)
> "Focus group interview was time-consuming." (Teacher group, Expert E4)

In summary, it could be inferred that in the analysis phase, teachers generally tended to employ direct observation for the initial assessment; laborious methods, such as interviews or questionnaires, were not recommended.

Applicability of identifying model factors with FDM

FDM was applicable to investigate the potential factors of the design model. This study could be considered an exploratory study because previous FDM research rarely explored the factors that constituted the design model. To examine validity, Delphi studies were frequently validated by follow-up research with interviews or surveys (Skulmoski et al., 2007). Accordingly, the qualitative data of the participating experts aligned with the FDM results, making the FDM a warranted technique to identify key factors. Furthermore, this study discovered that the FDM results also provided designers with detailed and customised references in an efficient way that was proved by the literature (Murray et al., 1985; Noorderhaben, 1995). Thus, FDM effectively facilitated developers to determine the priorities in each phase of the dynamic ADDIE model.

Conclusions and suggestions

This study aimed to identify the key factors and priorities associated with the development of instructional VR based on a dynamic ADDIE model. After two rounds of investigation, the “interactive feedback design” was selected as the optimal key factor, as both expert groups agreed on its improvement in student learning interest. In contrast, the “learner opinion survey” was ranked as the least influential factor. On the other hand, due to the result that the “formative evaluation” was the most essential sub-criterion, and the literature support, the dynamic ADDIE model established by this study was warranted. Moreover, owing to the reasons for divergence and preferences between expert groups, this study suggests that developers of different backgrounds should clearly identify the purpose of applying instructional VR, enabling them to choose alternatives that are suitable in the teaching context.

There are also some limitations in this study. First, since the number of school teachers developing VR in the country is scarce, the composition of the teacher group in this study was mostly elementary school teachers. Therefore, it is suggested that the teacher ratio can be adjusted more evenly. Second, although two rounds of FDM were performed, the reason for the divergence in the first three criteria of the teacher group remained ambiguous. Thus, it is recommended that future studies can conduct more interviews with in-service teachers.

Funding

This research was supported by the Higher Education Sprout Project of National Yang Ming Chiao Tung University (NYCU) and the Ministry of Education (MOE), Taiwan, and the National Science and Technology Council in Taiwan through Grant numbers MOST 111-2410-H-A49-018-MY4, MOST 110-2511-H-A49-009-MY2, and MOST 111-2423-H-A49-002.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Ruo-Yu Li is currently a Ph.D. student at the Institute of Education at the National Yang Ming Chiao Tung University. Her research interests comprise AR/VR/MR learning environments, neuroscience in educational technology, and the multiple-choice decision-making process.

Jerry Chih-Yuan Sun is a Distinguished Professor at the Institute of Education at the National Yang Ming Chiao Tung University where he leads the Interactive Learning Technology and Motivation Lab (see http://ILTM.lab.nycu.edu.tw). His research focuses on assessing innovative feedback technologies and their effects on student learning and motivation.

References

Akman, E., & Çakır, R. (2020). The effect of educational virtual reality game on primary school students' achievement and engagement in mathematics. *Interactive Learning Environments*, 1-18. https://doi.org/10.1080/10494820.2020.1841800

Allen, W. C. (2006). Overview and evolution of the ADDIE training system. *Advances in Developing Human Resources*, *8*(4), 430-441. https://doi.org/10.1177/1523422306292942

Alrehaili, E. A., & Al Osman, H. (2022). A virtual reality role-playing serious game for experiential learning. *Interactive Learning Environments*, *30*(5), 922-935. https://doi.org/10.1080/10494820.2019.1703008

Asad, M. M., Hassan, R. B., & Sherwani, E. F. (2014). Instructional models for enhancing the performance of students and workforce during educational training. *Academy Arena*, *6*(3), 27-31.

Asad, M. M., Sherwani, F., Hassan, R. B., Sahito, Z., & Khan, N. (2022). Workforce safety education and training for oil and gas industry: A conceptual framework for virtual reality-based HAZFO expert 2.0. *Journal of Engineering, Design and Technology*, *20*(5), 1122-1131. https://doi.org/10.1108/JEDT-08-2020-0330

Astuti, T. N., Sugiyarto, K. H., & Ikhsan, J. (2020). Effect of 3D visualization on students' critical thinking skills and scientific attitude in chemistry. *International Journal of Instruction*, *13*(1), 151-164.

Bogusevschi, D., Muntean, C. & Muntean, G.M. (2020). Teaching and learning physics using 3D virtual learning environment: A case study of combined virtual reality and virtual laboratory in secondary school. *Journal of Computers in Mathematics and Science Teaching, 39*(1), 5-18. https://www.learntechlib.org/primary/p/210965/

Branch, R. (2010). Instructional design: The ADDIE approach. Boston, MA: Springer US.

Buche, C., Bossard, C., Querrec, R., & Chevaillier, P. (2010). PEGASE: A generic and adaptable intelligent system for virtual reality learning environments. *International Journal of Virtual Reality*, *9*(2), 73-85. https://doi.org/10.20870/IJVR.2010.9.2.2772

Chang, C.-Y., Sung, H.-Y., Guo, J.-L., Chang, B.-Y., & Kuo, F.-R. (2022). Effects of spherical video-based virtual reality on nursing students' learning performance in childbirth education training. *Interactive Learning Environments*, *30*(3), 400-416. https://doi.org/10.1080/10494820.2019.1661854

Clark, R. C., & Mayer, R. E. (2016). *E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning.* John Wiley & Sons.

Cochrane, T. D., Cook, S., Aiello, S., Christie, D., Sinfield, D., Steagall, M., & Aguayo, C. (2017). A DBR framework for designing mobile virtual reality learning environments. *Australasian Journal of Educational Technology, 33*(6). https://doi.org/10.14742/ajet.3613

Dick, W., Carey, L., & Carey, J. O. (2009). *The Systematic Design of Instruction*. Merrill/Pearson.

Fiard, G., Selmi, S. Y., Promayon, E., Vadcard, L., Descotes, J. L., & Troccaz, J. (2014). Initial validation of a virtual-reality learning environment for prostate biopsies: Realism matters! *Journal of endourology*, *28*(4), 453-458. https://doi.org/10.1089/end.2013.0454

Fowler, C. (2015). Virtual reality and learning: Where is the pedagogy? *British journal of educational technology, 46*(2), 412-422. https://doi.org/10.1111/bjet.12135

Göksu, I., Özcan, K. V., Cakir, R., & Göktas, Y. (2017). Content analysis of research trends in instructional design models: 1999-2014. *Journal of Learning Design*, *10*(2), 85-109.

Gusmida, R., & Islami, N. (2017). The development of learning media for the kinetic theory of gases using the ADDIE model with augmented reality. *Journal of Educational Sciences*, *1*(1), 1-10.

Hu-Au, E., & Lee, J. J. (2017). Virtual reality in education: A tool for learning in the experience age. *International Journal of Innovation in Education*, *4*(4), 215-226.

Li, R.-Y., & Wang, C.-H. (2020). Key factors and network model for location-based cultural mobile game design. *British Journal of Educational Technology, 51*(6), 2495-2512. https://doi.org/10.1111/bjet.12926

Makransky, G., Andreasen, N. K., Baceviciute, S., & Mayer, R. E. (2021). Immersive virtual reality increases liking but not learning with a science simulation and generative learning strategies promote learning in immersive virtual reality. *Journal of Educational Psychology*, *113*(4), 719-735. https://doi.org/10.1037/edu0000473

Makransky, G., Petersen, G. B., & Klingenberg, S. (2020). Can an immersive virtual reality simulation increase students' interest and career aspirations in science? *British Journal of Educational Technology*, *51*(6), 2079-2097. https://doi.org/10.1111/bjet.12954

Murray, T. J., Pipino, L. L., & Van Gigch, J. P. (1985). A pilot study of fuzzy set modification of Delphi. *Human Systems Management*, *5*(1), 76-80. https://doi.org/10.3233/HSM-1985-5111

Noorderhaben, N. (1995). *Strategic decision making*. Addison-Wesley.

Parong, J., & Mayer, R. E. (2021). Cognitive and affective processes for learning science in immersive virtual reality. *Journal of Computer Assisted Learning*, *37*(1), 226-241. https://doi.org/10.1111/jcal.12482

Petersen, G.B., Klingenberg, S., Mayer, R.E. and Makransky, G. (2020). The virtual field trip: Investigating how to optimize immersive virtual learning in climate change education. *British Journal of Educational Technology, 51*(6), 2098-2114. https://doi.org/10.1111/bjet.12991

Sholihin, M., Sari, R. C., Yuniarti, N., & Ilyana, S. (2020). A new way of teaching business ethics: The evaluation of virtual reality-based learning media. *The International Journal of Management Education*, *18*(3), 100428. https://doi.org/10.1016/j.ijme.2020.100428

Skulmoski, G. J., Hartman, F. T., & Krahn, J. (2007). The Delphi method for graduate research. *Journal of Information Technology Education: Research*, *6*(1), 1-21. https://doi.org/10.28945/199

Smith, P. L., & Ragan, T. J. (2004). *Instructional design*. Wiley.

Soto, V. J. (2013). Which instructional design models are educators using to design virtual world instruction. *MERLOT Journal of Online Learning and Teaching*, *9*(3), 364-375.

Stavroulia, K. E., Christofi, M., Baka, E., Michael-Grigoriou, D., Magnenat-Thalmann, N., & Lanitis, A. (2019). Assessing the emotional impact of virtual reality-based teacher training. *The International Journal of Information and Learning Technology*, *36*(3), 192-217.

Trevelyan, E. G., & Robinson, N. (2015). Delphi methodology in health research: How to do it? *European Journal of Integrative Medicine*, *7*(4), 423-428. https://doi.org/10.1016/j.eujim.2015.07.002

Trust, T., & Pektas, E. (2018). Using the ADDIE model and universal design for learning principles to develop an open online course for teacher professional development. *Journal of Digital Learning in Teacher Education*, *34*(4), 219-233. https://doi.org/10.1080/21532974.2018.1494521

Wang, P., Wu, P., Wang, J., Chi, H. L., & Wang, X. (2018). A critical review of the use of virtual reality in construction engineering education and training. *International journal of environmental research and public health*, *15*(6), 1204. https://www.mdpi.com/1660-4601/15/6/1204

Warren, S. J., Lee, J., & Najmi, A. (2014). The impact of technology and theory on instructional design since 2000. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology* (pp. 89-99). Springer New York. https://doi.org/10.1007/978-1-4614-3185-5_8

Wei, W.-L., & Chang, W.-C. (2008). Analytic network process-based model for selecting an optimal product design solution with zero-one goal programming. *Journal of Engineering Design*, *19*(1), 15–44. https://doi.org/10.1080/09544820601186054

Wiphasith, H., Narumol, R., & Sumalee, C. (2016). The design of the contents of an e-learning for teaching M. 5 English language using ADDIE model. *International Journal of Information and Education Technology*, *6*(2), 127.

Yu, S.-J., Hsueh, Y.-L., Sun, J. C.-Y., & Liu, H.-Z. (2021). Developing an intelligent virtual reality interactive system based on the ADDIE model for learning pour-over coffee brewing. *Computers and Education: Artificial Intelligence*, *2*, 100030. https://doi.org/10.1016/j.caeai.2021.100030