

The Construction of the Evaluation Index System of Children's Educational Game Learning Accessibility

Xueying Yao¹, Tianyu Zhou², Yafeng Niu³

*¹ School of Mechanical Engineering, Southeast University
Nanjing, China*

ABSTRACT

The quality of children's educational games affects children's cognitive development and skills. Therefore, the construction of educational game evaluation indexes is playing the big role in the development of educational games. In this paper, the initial evaluation indexes of educational game learning accessibility were developed by applying the relevant theories. Then three rounds of expert evaluations were conducted using the Delphi Method, and the initially formulated educational game learning accessibility evaluation indexes were revised and improved according to the experts' suggestions. This paper determines the educational game learning accessibility evaluation indexes including 3 primary indicators, 7 secondary indicators and 26 tertiary indicators. Finally, two-by-two comparison of indicators at each level is conducted to determine the weight of each evaluation indicator by means of analytic hierarchy process. This evaluation index system can facilitate children's parents and teachers to select educational games according to the indexes, and also provide an effective reference basis for the design, improvement and evaluation of children's educational games in the future.

Keywords: Children's educational games, Evaluation index, Accessibility

INTRODUCTION

When evaluating educational games, different evaluation criteria should be used according to the different stages of the population and different educational purposes. The establishment of the evaluation indexes of educational games should be a joint effort of users, game design experts, and experts in the field of education. As the digitalization process accelerates and educational games are constantly introduced, their evaluation index system will be constantly revised and improved to meet the changes of the times (Jeon et al. 2017; Wu and Chen, 2013). Therefore, we will build a system of evaluation indexes for the accessibility of educational games based on the research and suggestions of experts in various fields.

PRELIMINARY ESTABLISHMENT INDEX

Education is a result of the interaction of formation and process, not a one-way transmission. The accessibility of learning is not only reflected in the performance of educational results, but also includes the basic condition of the learner and a series of physiological and psychological reactions produced in the learning process. According to the process of education, the indicators of educational game learning accessibility evaluation are divided into three primary indicators: basic, process and summative evaluation (Dick, 1977).

Basic evaluation

The main subjects of this study were school-age children (6-7 years old). According to seven-dimensional evaluation scale of gamified online learning (De Freitas and Oliver, 2006), we focus on the learners' profile, such as their age, knowledge background and interests. Meanwhile, Fan takes learners' characteristics, information literacy, internal and external motivation and experience as the main evaluation indicators (Fan and Cui, 2008). Firstly, the basic evaluation should evaluate children with full consideration of different children's physiological and psychological characteristics as well as their personal wishes. Secondly, from the perspective of game design, we will consider the game design itself in terms of its richness, rationality and aesthetic value from the dimensions of playfulness, technicality and artistry (Ye et al. 2009).

Process evaluation

Process evaluation is the evaluation of children's performance, their personal feelings and attitudes while using educational games. According to the User Experience Evaluation of Educational Games (UEREK), this index system is based on the user experience perspective, involving the user's perceived aesthetics, usability, needs and so on (Shen, 2015).

Summative evaluation

Summative evaluation serves as a judgment of the degree of educative goals of educational games. It is also a measure of the adaptability of educational goals and effectiveness of strategies. It is usually carried out by means of quantitative evaluation such as exams and tests.

Preliminary determination of learning evaluation indexes

There are the initial evaluation indexes of educational game learning accessibility developed based on the literature analysis (see Table 1).

Table 1: Educational Game Learning Accessibility Evaluation Index (Version 1)

Tier 1 Indicators	Secondary indicators	Tertiary indicators	Description	
Basic evaluation	Learner Dimension	Learner Information	Age, gender, education level of learners	
		Learner Willingness	Learners' willingness to use the game, hobbies	
	Game Design Dimension	Subject Type	Selection of applicable disciplines	
		Course Compliance	The content of the game fits the content of the textbook materials	
		Content Difficulty	Level of difficulty of the content of the topic	
		Feedback Reasonableness	Positive, effective and encouraging feedback mechanisms	
Process evaluation	Form preference	Interaction way fun	The difference in the fun of using different interaction methods	
		Interaction style ease of use	Differences in ease of use due to different interaction methods	
		Sense of accomplishment	Differences in inner sense of accomplishment when answering questions by different interaction methods	
	Achievement Impact	Concentration	The effect of different interaction methods on the concentration level during the game	
		Memory	The effect of different interaction methods on the memory of game answer knowledge	
		Completion	The effect of different interaction methods on adherence to the end of the game	
	Comfort of use	Pressure level	The effect of different interaction methods on user stress	
		Tension level	The effect of different interaction methods on the user's tension level when using	
		Naturalness	Naturalness when using different interaction methods	
		Degree of accuracy	The effect of different interaction methods on the accuracy of answer choices	
	Summative evaluation	Knowledge Mastery	Maintain test results	Use tests and other methods to test what you have learned
		Knowledge transferability	Migration test results	Use tests and other methods to design new questions to test the changes in knowledge learned

INDEX OPTIMIZATION

The Delphi method is a qualitative assessment method that adds, removes and modifies the selected indicators based on the experience and recommendations of the experts (Liu et al. 2011). After no less than three rounds of cyclical interaction process with experts, it allows to reach a coherent result for the opinions.

Questionnaire distribution and analysis

The indicators are described in the questionnaire. We hope that experts will suggest and propose modifications to the unreasonable indicators. The questionnaires were distributed to 15 senior teachers in Primary School and 5 senior designers in the educational game industry for a total of 20 people. The questionnaires were distributed again after modification according to the experts' suggestions for a total of 3 rounds. The positive coefficients of experts in the three rounds of questionnaires were 100%, 80%, and 75%, respectively, to meet the needs of the study.

According to the results and opinions of experts' scoring of indicators, the relevant indicators were modified or deleted. After three rounds of experts' questionnaires, it was found that the average value of the scores for each indicator of the educational game learning accessibility evaluation indicators increased, with the average value rising from 3.743 to 4.590, indicating the experts' approval of the evaluation indicators after two modifications.

According to the three rounds of questionnaire, the coefficient of variation did not exceed 0.3, and the coefficient of variation gradually decreased, indicating that the experts' opinions reached a high degree of agreement. The coordination coefficient of experts' scoring gradually increased in the three rounds of the questionnaire, and reached 0.186 in the third round of the questionnaire. The p-values were all less than 0.05, indicating that the experts' scoring had a high degree of consistency.

There are some problems with the initially developed indicators for evaluating the accessibility of educational game learning.

From the perspective of basic evaluation, the learners' dimension can be increased by the indicators of operating environment. A familiar and quiet environment determines the state of learners. The supporting adaptive equipment is also the basic preparation for playing games. Learners' past experiences also reflect the difference of their own adaptive ability. The indicators of interface design in the game design dimension should also be enriched. Appropriate external incentives and guidance in the operation process can make the learners more motivated when using the educational games.

From the perspective of process evaluation, experts take into account the Mihaly's FLOW theory, adding the state of immersion to it. It indicates that learners are able

to fully engage in the situation during educational games and may drive themselves to complete tasks that they would not normally be able to do (Wang et al. 2021).

From the perspective of summative evaluation, experts believe that the knowledge mastery and knowledge transferability degree in the secondary indicators are the same indicator, so they should be combined into knowledge mastery degree. The review and reflection of the whole learning process should also be taken into consideration.

Finally, there are the final results of the indicators (see Table 2).

Table 2: Educational game learning accessibility evaluation indicators (final version)

Tier 1 Indicators	Secondary indicators	Tertiary indicators	Description
Basic evaluation	Learner Dimension	Learner Information	Age, gender, education level of learners
		Learner Willingness	Learners' willingness to use the game, hobbies
		Operating Environment	The environment and equipment in which learners use the game
		Game Learning Experience	Used educational game experience, genre, style
	Game Design Dimension	Subject Type	Selection of applicable disciplines
		Course Fit	The content of the game fits the content of the textbook materials
		Content Difficulty	Level of difficulty of the content of the topic
		Feedback Reasonableness	Positive, effective and encouraging feedback mechanisms include action feedback and answer feedback
		Incentive Mechanism	External incentives such as points and flowers
		Interface Friendliness	Interface meets the use of special groups of children
		Social Value	Positive value system, no sensitive topics
		Operation guidance	Provide guidance and answer questions that arise during use
	Interface artistry	Beautiful interface, reasonable color matching, good visual effect	
Process evaluation	Form preference	Interaction way fun	The difference in the fun of using different interaction methods
		Interaction style ease of use	Differences in ease of use due to different interaction methods
		Sense of accomplishment	Differences in inner sense of accomplishment when answering questions by different interaction methods
		Immersion	The difference in immersion during the game brought by different interaction methods
	Achievement Impact	Concentration	The effect of different interaction methods on the concentration level during the game
		Memory	The effect of different interaction methods on the memory of game answer knowledge
		Completion	The effect of different interaction methods on adherence to the end of the game
	Comfort of use	Pressure level	The effect of different interaction methods on user stress and tension generation
Naturalness		Naturalness when using different interaction methods	

		Degree of accuracy	The effect of different interaction methods on the accuracy of answer choices
Summative evaluation	Knowledge Mastery	Retention test	Use paper and pencil tests and other methods to test what you have learned
		Transfer test	Use paper and pencil tests and other methods to design new questions to test the changes in knowledge learned
	Learning to be reflective	Learning Summary	Reflecting on the whole learning process in the form of a summary

THE WEIGHTS OF THE EVALUATION INDEXES

This part combines the Delphi Method and the analytic hierarchy process (Zhai et al. 2019), and generates a judgment matrix by issuing questionnaires and comparing all primary, secondary and tertiary indexes two by two. By calculating analysis of the weight of each indicator. In the questionnaire, a matrix scale of 1-9 was used, from extremely unimportant 1 to extremely important 9, where 5 represents equal importance. A total of 20 questionnaires were distributed to teachers and designers. A total of 16 questionnaires with a CR < 0.1 (valid questionnaires) were tested for consistency, which meets the requirement of no less than 10 expert evaluations.

Calculation of index weights

By group decision making, the expert data set method is a weighted arithmetic average of the individual expert ranking vectors. There are the final group decision calculation results (see Tables 3 and 4).

Table 3: Expert weighted arithmetic average weights (Tertiary Indicators)

	Tertiary indicators	Weights		Tertiary indicators	Weights
1	Learner Information	0.2160	14	Completion	0.0131
2	Learner Willingness	0.1486	15	Content Difficulty	0.0127
3	Interaction way fun	0.1071	16	Immersion	0.0122
4	Concentration	0.0963	17	Learning Summary	0.0121
5	Interaction style ease of use	0.0699	18	Naturalness	0.0116
6	Hold test	0.0586	19	Migration Testing	0.0108
7	Memory	0.0574	20	Feedback Reasonableness	0.0087
8	Operating Environment	0.0406	21	Incentive Mechanism	0.0057
9	Game Learning Experience	0.0269	22	Interface Friendliness	0.0039
10	Sense of accomplishment	0.0221	23	Degree of accuracy	0.0028
11	Subject Type	0.0212	24	Social Value	0.0026
12	Pressure level	0.0182	25	Operation guidance	0.0019
13	Course Fit	0.0178	26	Interface artistry	0.0014

Table 4: Expert weighted arithmetic mean weights (Secondary & Tier 1 Indicators)

Secondary indicators	Weights	Tier 1 Indicators	Weights
Learner Dimension	0.4321	Basic evaluation	0.5079
Form preference	0.2113	Process evaluation	0.4106
Achievement Impact	0.1668	Summative evaluation	0.0815
Game Design Dimension	0.0758		
Knowledge Mastery	0.0694		
Comfort of use	0.0325		
Learning to be reflective	0.0121		

Discussion

Basic evaluation is the most important indicator of the accessibility of educational games. The user's physical state, personal information and subjective intentions are all prerequisites for the design and development of an educational game as well as for the choice of educational genre. The game design dimension is less influential than the learner dimension in terms of accessibility, but is also in the middle of the range, suggesting that reasonable feedback, user-friendly interface design and other relevant indicators are also worthy of attention.

Process evaluation is second to basic evaluation. Experts tend to focus on the experience of children when using. Both the innovative interaction and the development of learning habits are intended to stimulate children's interest and curiosity.

The summative evaluation has a relatively low weighting in the first level of indicators. This suggests that we should avoid the psychological burden of too much education on children, which reduces the attractiveness of the educational game itself.

CONCLUSIONS

This paper combines the methods of educational evaluation to ensure that the evaluation index of educational game learning accessibility has a relatively mature theoretical foundation. From the 3 tertiary indicators of basic evaluation, process evaluation and summative evaluation, a total of 7 secondary indicators have been constructed, as well as 26 tertiary indicators. Through the calculation of the weight of each index, it reflects the focus direction and research and development strategies of experts in education and game related fields at the present stage. It also reflects that the evaluation index has certain feasibility and scientificity. If the requirements of the above indicators are met, an educational game suitable for children can be designed.

ACKNOWLEDGMENTS

This work was supported jointly by National Natural Science Foundation of China [No. 72171044, 71801037], Aerospace Science Foundation of China [No. 20200058069002], Southeast University "Zhi Shan Qing Nian Xue Zhe" Support program funding.

REFERENCES

- De Freitas, S., & Oliver, M. (2006). How can exploratory learning with games and simulations within the curriculum be most effectively evaluated?. *Computers & education*, 46(3), pp.249-264.
- Dick W. (1977). Summative evaluation. *Instructional design: Principles and applications*, pp.337-348.
- Fan, Y.H., Cui, J.Y. (2008). Research on the development and application of online educational game evaluation gauges. *The Chinese Journal of ICT in Education*, (06), pp.10-12.
- Jeon, J., Yoon, D., Yang, S., & Kim, K. (2017, August). Extracting gamers' cognitive psychological features and improving performance of churn prediction from mobile games. *2017 IEEE Conference on Computational Intelligence and Games (CIG)*, pp.150-153.
- Liu, W. T., Gu, H., & Li, C. H. (2011). Expert evaluation method based on Delphi method. *Computer Engineering*, 37(S1), pp.189-191+.
- Shen, J. (2015). *A comparative research on evaluational index of educational games in China and the United States*. Hangzhou Normal University.
- Wang, S., Yin, Y., Wang, T., Luo, J. L. (2021). Flow experience in educational context. *Journal of Educational Biology*, 9(01), pp.59-64.
- Wu, Y., & Chen, V. H. H. (2013). A social-cognitive approach to online game cheating. *Computers in Human Behavior*, 29(6), pp.2557-2567.
- Ye, C.Q., Wang H.Y., Wang P. (2009) A three-dimensional evaluation system architecture for digital teaching games. *Journal of Distance Education*, 17(06), pp.71-73.
- Zhai, J., Wang, X.J., Zhao, G.Y., Hou, Y., Wang, B.S. (2019). Construction of evaluation index system of medical magnetic resonance imaging equipment experience design based on Delphi method and hierarchical analysis. *China Medical Devices*, 34(8), pp.11-14.