

Quality Measurement for Serious Games

Louis Khrisna Putera Suryapranata
Doctor of Computer Science
Bina Nusantara University
Jakarta, Indonesia
louis.suryapranata@binus.ac.id

Benfano Soewito
Master in Computer Science
Bina Nusantara University
Jakarta, Indonesia
bsoewito@binus.edu

Gede Putra Kusuma
Master in Computer Science
Bina Nusantara University
Jakarta, Indonesia
inegara@binus.edu

Ford Lumban Gaol
Doctor of Computer Science
Bina Nusantara University
Jakarta, Indonesia
f.gaol@binus.edu

Harco Leslie Hendric Spits Warnars
Doctor of Computer Science
Bina Nusantara University
Jakarta, Indonesia
shendric@binus.edu

Abstract—One kind of games based on its playing media is video game, which defined as a game that use any kind of computers as media. In terms of purpose, game has two types, which is serious game and non-serious game. As learning media, serious games need to have ability to motivate its player in order to play the game until the end so the player can finish the game while understanding learning materials given in the game. Therefore, the quality of serious game should be measured, like any other software. When it comes to measure serious games, previous models can't do the job well since all of them have no way to measure the game's content, especially to measure the content that can improve player's motivation. This paper proposes a new game measurement metric to measure the game quality factors, especially to measure serious game contents that can motivate its players. The result of this paper is a metric table consisted of measureable factors and type of test to measure them.

Keywords—quality measurement; quality metric; serious games; quality factors; motivational factors

I. INTRODUCTION

Game is one kind of play activity, in a pretended reality, which player(s) try to achieve at least one goal by acting based on rules [3]. Game has three elements, which are play activity, fun, and entertainment [15]. Game can be applied to any aspect of life and give benefits, such as increased motivation in learning subjects given in classes, better and faster performance in surgical operations, enhanced problem-solving skills, and increased cognitive flexibility [5]. Game could be used as learning media because game could increase students' self-efficacy because they were immersed and engaged with the game, while they didn't recognize that they have learnt something given in the game [11]. Moreover, Virtual Information System can be applied in order to present playing and having fun atmosphere on working area [16]. Virtual application can be built in order to run and support intelligent application in virtual information systems with technologies such as Data warehouse [17,18] or Data mining such as Attribute Oriented Induction algorithm [19,20] or AOI-HEP algorithm with ability to find frequent and similar patterns [21,22,23].

One kind of games based on its playing media is video game, which defined as a game that use any kind of computers as media. Like any kind of software, game should be measured to know about its quality [13]. Software quality measurement metric can't be used to measure game quality because game is different compared to other type of application. The difference is, game is consisted from scenarios that built by combining visual aspect and audio aspect in order to make player enjoy them [1], thus we need another way to measure game quality. [1], [13], [14] have made some quality measurement standards by modifying software measurement metrics based on several quality criteria such as ISO 25010, IEEE STD-1061, ISO 25022, and ISO 9126-3 to measure game's quality, divided to several aspects: usability, flexibility, content, gameplay, fun, and so on.

In terms of purpose, game has two types, which is serious game and nonserious game. The difference between serious games and nonserious games is the focus, while non-serious games focus on pure entertainment, serious games focus on educational purposes [2]. Serious game is a game that support the player to achieve learning targets while having fun in the process [12]. Serious game can be used as learning media because of several reasons, such as the ability to support constructive, experiential, situated, and procedural learning; engage and motivate players, and promote self-regulated learning [7]. As learning media, serious games need to have ability to motivate its player in order to play the game until the end so the player can finish the game while understanding learning materials given in the game. When it comes to measure serious games, previous models can't do the job well since all of them have no way to measure the game's content, especially to measure the content that can improve player's motivation.

This paper proposes a new game measurement metric to measure the game quality factors and type of measurement instruments that appropriate to each factor, especially to measure serious game contents that can motivate its players. The subsequent sections will cover as follows: the related works in section 2, process of defining metrics and the result in section 3, conclusion and future work in the last section.

II. RELATED WORKS

A. *Serious Game*

The advancement of technology had some impacts that influenced the development of game technology, such as augmented reality, virtual reality, and serious games [4]. Serious game is a game that support the player to achieve learning targets while having fun in the process [12]. The fun aspects can be determined by several factors, such as storyboard, usability, graphics, mechanisms, and devices involved to play. The difference between serious games and nonserious games is the focus, while non-serious games focus on pure entertainment, serious games focus on educational purposes [2].

Although, at first, serious game gets less attention as research topic [8], it presents a promising opportunity as learning media [6]. Serious game can be used as learning media because of several reasons, such as the ability to support constructive, experiential, situated, and procedural learning; engage and motivate players, and promote self-regulated learning [7].

B. *Game Quality Measurement*

[14] wrote about the measurement of mobile game's usability quality. Using GQM method, they propose a metric table consist of attributes of usability quality that would be measured, questions related to those attributes, things related to the attributes that would be measured, and the relation of those attributes to ISO 9126-3, along with how to measure those attributes. The metric consist of three categories, which are content, device, and gameplay; based on five sub-attributes of usability quality, which are understandability, learnability, operability, attractiveness, and compliance. Understandability is related to player's perception of how to control their characters and how the rules work in the game. Learnability is related to player's ability to play the game after watching tutorial. Operability is related to player's ability to custom the game settings and content according to their preferences. Attractiveness is related to the appeal of user interface and gameplay method based on player's preference. Compliance is related to difficulty of rules applied in the game, either rules for gameplay or device requirement. They also write steps to apply the proposed metric in measurement process. From their study, it can be concluded this metric can be used to measure the usability quality of any kind of games, especially mobile games.

[1] wrote about the measurement of game's quality. They propose a metric table consist of quality and sub quality related to ISO 25010, metric name and description, evaluation method, and relation of the factor to ISO 25022. The metric consist of three categories, which are usability (with sub-quality effectiveness, efficiency, and satisfaction), flexibility (with sub-quality accessibility), and safety (with sub-quality user health and safety). Effectiveness is related to player's performance in order to achieve the goal in the game. Efficiency is related to player's performance in using resources provided in game to achieve the goal. Satisfaction is related to player's feeling while playing the game. Accessibility is related to the player's ability to play the game in various ways. User

health and safety is related to the effect that might happened on the player while playing the game either physically or mentally. They use the metric to evaluate 5 mobile games from different game developer, different genre, and different platform. From their study, it can be concluded this metric can be used to measure the quality of any games.

[13] wrote about development of game testing method to measure the quality of the game. They propose a method to measure the game's quality by combining existing game testing approach with software testing approach, and the criteria measured are the result of combining software engineering quality and game quality, which is a combination from ISO 25010, IEEE STD-1061, and Fullerton Model. Their proposed testing method is used to test several factors, which are user experience (with sub factors fun, balance, and usability), functional (with sub factors feature availability, performance, internal completeness, and service compatibility), maintainability, and portability. Fun is related to the engagement of player while playing the game. Balance is related to difficulty level of the game and the fairness of reward and punishment available in the game. Usability is related to player's experience when facing the game's user interface. Feature ability is related to the completeness of game feature to work as expected. Performance is related to the usage of resources while the game runs. Internal completeness is related to the completeness of game rules so the game can have full functionality. Service compatibility is related to the completeness of external service to support the game. Maintainability is related to the easiness of game code maintenance. Portability is related to the easiness of transferring the game code from the current platform to other platforms. Their proposed test method divided to three steps, which are pre-playtest, playtest, and post-playtest. Each step focus on different factors and using different test methods. From their study, it can be concluded that their proposed method can be used to measure game quality and provide some useful information to be generated in game quality report.

C. *Motivation in Learning*

Motivation is a problem for both teachers and students [10]. To have learners motivated, there are several aspects that must be accomplished. First, they must be curious about topic given. Second, the instruction must be relevant for them to reach the goal. Third, they must feel confident that they can reach their goal. Last, the result of learning should make learners satisfied. Those aspects above are the base of motivational measurement model called ARCS [9]. ARCS divides motivation of learners into 4 aspects, which are attention, relevance, confidence, and satisfaction. Attention is related to how much the learning material can attract the learner. Relevance is related to the relation between the learning material with things learner faced in their life. Confidence is related to the ability of learning material to boost learner's confidence after completing the learning process. Satisfaction is related to learner's feeling while and after completing the learning process. ARCS can also be used to measure how much learning media can motivate learners to learn something [10].

III. QUALITY METRIC FOR SERIOUS GAMES

Before we proposed a quality measurement for serious game, first, we had to analyze the software quality factors relevant to this research. After factors analysis, all factors are grouped based on two categories, which is player related and software related. Aspects that can be seen and felt by players are categorized in player related group, while aspects that can't be seen and felt by players, but can be checked by other computer programs are categorized in software related group.

Components of ARCS are integrated into the grouped factors to add motivational factor into our proposed metric. Attention is related to how much the game can attract the player. Relevance is related to the relation between the content inside the game with things player faced in their life. Confidence is related to the ability of the game to boost player's confidence after completing the game. Satisfaction is related to player's feeling while and after completing the game. Integration result can be seen in Table 1.

After integrating ARCS with quality factors, those factors are simplified to reduce the redundancy among factors. Fun, attractiveness, and compliance are grouped into satisfaction category, since all of them are related to player's satisfaction while and after playing the game.

TABLE I. QUALITY CATEGORY GROUPING

Category Groups	Quality Factors			
	Ramadan & Hendradjaya [13]	Pavapootanont & Prompoon [14]	Trisnadoli et al. [1]	Keller [10]
Player related	Fun	Understandability	Effectiveness	Attention
	Balance	Attractiveness	Efficiency	Relevance
	Usability	Learnability	Satisfaction	Confidence
		Compliance	User health and safety	Satisfaction
		Operability	Accessibility	
Software related	Performance			
	Internal completeness			
	Service compatibility			
	Feature ability			
	Maintainability			
	Portability			

Understandability, learnability, operability, accessibility, and user health and safety are grouped into usability category, since all of them are related to user experience while playing with the game. Effectiveness and efficiency are grouped into balance category, since they are related to game balance. To complete our proposed metric, we also include type of measurement instruments that suited for each category. Player performance can be seen from player's progress in the game, player's character level, player's highest score, player's win/lose rate, and so on, depends on the kind of game. Higher

indicator value indicates better player performance. Player questionnaire consists of questions to get player's opinion about balance of rules in the game, usability experience, and player's motivational aspect (which separated into attention, relevance, confidence, and satisfaction) which using Likert scale from 1 to 5 with 1 as very bad, 2 as bad, 3 as medium, 4 as good, and 5 as very good. Higher mean score indicates better quality for scored factor. Application performance can be concluded from player report about glitches, bugs, and crashes while playing with the game as indicator of game performance, internal completeness, feature ability, and service compatibility. Player report are classified into four types, and the amount of report scoring system for each category; 0 report as very good, 1-3 reports as good, 4-6 reports as medium, 7-9 reports as bad, and 10 and up reports as very bad. Developer team questionnaire consisted of questions to get developer's opinion about service compatibility, maintainability, and portability which using Likert scale from 1 to 5 with 1 as very bad, 2 as bad, 3 as medium, 4 as good, and 5 as very good. Higher mean score indicates better quality for scored factor. Our proposed quality measurement metric can be seen in Table 2.

TABLE II. PROPOSED QUALITY MEASUREMENT METRIC

Category Groups	Quality Factors	Measurement Instruments
Player related	Balance	Player performance & player questionnaire
	Usability	
	Attention	Player questionnaire
	Relevance	
	Confidence	
Satisfaction		
Software related	Performance	Application performance
	Internal completeness	
	Feature ability	
	Service compatibility	
	Maintainability	Developer team questionnaire
	Portability	

IV. CONCLUSION & FUTURE WORK

In this paper, we proposed a new game measurement metric to measure the game quality factors, especially to measure serious game contents that can motivate its players. We found that previous quality measurement metrics can't measure the quality of serious game because all of them haven't stated the way to measure the motivational aspect, which is important for a serious game to be able to motivate its player to play the game until the end. Our proposed metric has included how to measure motivational factors based on ARCS, which previous proposed metric don't have.

This paper has several limitations, such as restricted criteria of papers used as information source, which only papers related to game quality measurement that used as references. Another

limitation is the metric proposed in this paper haven't been used to test the quality of serious games. For our future work, we will use papers related to software quality measurement to complete metric proposed in this paper, since game is one kind of software, then game might have similarities with nongame applications. We also plan to use the metric proposed in this paper to test the quality of several serious games in our next study.

REFERENCES

- [1] A. Trisnadoli, B. Hendradjaya, and W. D. Sunindyo, "A Proposal of Quality Model for Mobile Games," 5th Int. Conf. Electr. Eng. Informatics 2015, pp. 377–381, 2015.
- [2] D. Djaouti, J. Alvarez, J.-P. Jessel, and O. Rampoux, "Origins of Serious Games," *Serious Games Edutainment Appl.*, pp. 9–23, 2011.
- [3] E. Adams, *Fundamentals of Game Design*, 2nd ed. CA: New Riders, 2009.
- [4] E. F. Anderson, L. McLoughlin, F. Liarokapis, C. Peters, P. Petridis, and S. de Freitas, "Developing serious games for cultural heritage: A state-of-the-art Review," *Virtual Real.*, vol. 14, no. 4, pp. 255–275, 2010.
- [5] Entertainment Software Association, *Games Improving Education*. 2014.
- [6] F. Bellotti, R. Berta, A. De Gloria, A. D'ursi, and V. Fiore, "A serious game model for cultural heritage," *J. Comput. Cult. Herit.*, vol. 5, no. 4, pp. 1–27, 2012.
- [7] F. M. Dagnino et al., "Serious Games to Support Learning of Rare 'Intangible' Cultural Expressions," 9th Int. Technol. Educ. Dev. Conf. (INTED 2015), no. March, pp. 7184–7194, 2015.
- [8] J. Froschauer, M. Arends, D. Goldfarb, and D. Merkl, "A serious heritage game for art history: Design and evaluation of ThIATRO," in *Proceedings of the 2012 18th International Conference on Virtual Systems and Multimedia, VSMM 2012: Virtual Systems in the Information Society*, 2012, no. September 2012, pp. 283–290.
- [9] J. M. Keller, "First principles of motivation to learn and e 3 -learning," *Distance Educ.*, vol. 29, no. 2, pp. 175–185, 2008.
- [10] J. M. Keller, "Motivation, Learning, and Technology: Applying the ARCS-V Motivation Model," *Particip. Educ. Res.*, vol. 3, no. 2, pp. 1–13, 2016.
- [11] L. Yin, L. Ring, and T. Bickmore, "Using an interactive visual novel to promote patient empowerment through engagement," *Proc. Int. Conf. Found. Digit. Games - FDG '12*, no. May, pp. 41–48, 2012.
- [12] M. Mortara, C. E. Catalano, F. Bellotti, G. Fiucci, M. Houry-Panchetti, and P. Petridis, "Learning cultural heritage by serious games," *J. Cult. Herit.*, vol. 15, no. 3, pp. 318–325, 2014.
- [13] R. Ramadan and B. Hendradjaya, "Development of game testing method for measuring game quality," in *Proceedings of 2014 International Conference on Data and Software Engineering, ICODSE 2014*, 2015.
- [14] S. Pavapootanont and N. Prompoon, "Defining usability quality metric for mobile game prototype using software attributes," in *Proceedings of the IEEE International Conference on Software Engineering and Service Sciences, ICSESS*, 2015, vol. 2015–Novem, pp. 730–736.
- [15] H.L.H.S Warnars, "Game Information System," *Int. J. Comput. Sci. Inf. Technol.*, vol. 2, no. 3, pp. 135–148, 2010.
- [16] H.L.H.S Warnars, (2008), "Virtual Information system on Working area", Indonesian Students' International Scientific Meeting, (Temu Ilmiah Internasional Mahasiswa Indonesia, TIIMI), pp. 169-181, London, United Kingdom, 5-7 December 2008. Source : <http://arxiv.org/ftp/arxiv/papers/1006/1006.1701.pdf>
- [17] H.L.H.S Warnars, R. Randriatomanana (2016), "Datawarehouse: A Data Warehouse artist who have ability to understand data warehouse schema pictures", IEEE TENCON 2016 (Technologies for Smart Nation), pp. 2207-2210, 22-25 Nov 2016, Singapore. <http://site.tencon2016.focalevents.sg/>
- [18] H.L.H.S Warnars, (2014), "Perbandingan penggunaan Database OLTP (Online Transactional Processing) dan Data Warehouse", *Creative Communication and Innovative Technology (CCIT) journal*, Vol. 8 No. 1, pp. 83-100, September 2014. ISSN: 1978-8282.
- [19] H.L.H.S Warnars (2015), "Mining Patterns with Attribute Oriented Induction", *The International Conference on Database, Data Warehouse, Data Mining and Big Data (DDDDBD2015)*, Tangerang, Indonesia, pp. 11-21, 10-12 September 2015.
- [20] H.L.H.S Warnars, M I Wijaya, H B Tjung, D F Xaverius, D V Hauten, Sasmoko (2016), "Easy understanding of Attribute Oriented Induction (AOI) characteristic rule algorithm", *International journal of Applied Engineering Research (IJAER)*, vol. 11, No. 8, pp. 5369-5375. ISSN-0973-4562. Source: http://www.ripublication.com/ijaer16/ijaerv11n8_03.pdf
- [21] H.L.H.S Warnars (2016), "Using Attribute Oriented Induction High level Emerging Pattern (AOI-HEP) to mine frequent patterns", *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 6, No.6 <http://www.iaesjournal.com/online/index.php/IJECE/article/view/10579>
- [22] H.L.H.S Warnars, (2014), "Mining Frequent and Similar Patterns with Attribute Oriented Induction High Level Emerging Pattern (AOI-HEP) Data Mining Technique", *International Journal of Emerging Technologies in Computational and Applied Sciences (IJETCAS)*, vol. 3, Issue 11, pp.266-276. EISSN : 2279-0047, ISSN: 2279-0055. Source : <http://iasir.net/IJETCASpapers/IJETCAS15-247.pdf>
- [23] H.L.H.S Warnars (2014), "Mining Frequent Pattern with Attribute Oriented Induction High level Emerging Pattern (AOI-HEP)", *IEEE the 2nd International Conference on Information and Communication Technology (IEEE ICoICT 2014)*, Bandung, Indonesia, pp. 144-149, 28-30 May 2014.