

A Serious Game for Improving Reading Skill of Preschoolers

Muhammad Rizki Nugraha¹, Yusep Rosmansyah^{2*}, Evi Dewi Sri Mulyani³, Teuku Mufizar⁴,
Dede Syahrul Anwar⁵, Cepi Rahmat Hidayat⁶

^{1,3,4,5,6}Informatics STMIK Tasikmalaya

²Informatics Institut Teknologi Bandung

*E-mail: yusep@stei.itb.ac.id

ABSTRACT

Nowadays, many games are made for educational or learning purposes. However, the unresolved challenge remains, and the perfect balance between the fun and learning aspects remains. This paper represents an effort toward such a challenge by way of game difficulty dynamization implementation in a game to improve preschoolers reading skills using Dynamic Difficulty Adjustment dan Procedural Content Generation. To measure the balance of fun and learning aspects in the game, experiments were conducted involving 32 preschoolers. The proposed game can generate 10.962 game content in various difficulties and balance the fun and learning aspects according to the enjoyment factor's average value. Students could enjoy the game while significantly gaining an increase in the average reading skill, based on the t-test value and NGain value of 0.551.

Keywords: serious game, dynamic difficulty adjustment, procedural content generation, reading skills.

INTRODUCTION

Reading means seeing and understanding the contents of what is written, spelling, or reciting what is written or pronounced [1]. The ability to read is a basic ability that someone needs for the need to communicate and interact. However, the Indonesian statistics department (BPS) explains that in Indonesia up to 2018, around 2.07% of the Indonesian population is still illiterate with an age range of 15-59 years [2]. On the other hand, smartphone users are increasing. Based on data from E-Marketers and Hootsuite, smartphone users in Indonesia up to 2018 have reached around 170 million [3], [4]. As smartphone usage increases, the number of mobile gamers also increases. Based on a survey conducted by the Decision Lab and Mobile Marketing Association it is known that the number of mobile gamers in Indonesia in 2018 reached 60 million [5]. The survey also said that 25% of the activity on smartphones is playing games for around 53 minutes [5]. 72% of the 3917 children aged 3-8 years are mobile gamers with an average duration of playing more than 60 minutes each day [6]. Based on these data, games can be used as a medium to improve reading

skills. However, according to Go-Globe and WePC, the education game genre is less desirable and only has a popularity percentage of 8.89% [7], [8]. Therefore, a method to make the game more interesting is needed.

Games made for educational or learning purposes are serious games. According to Michael and Chen, serious games can be interpreted as games with an element of learning so that the details of games such as entertainment and pleasure are no longer the main focus [9]. In previous research, Rambli [10] and Ferrer [11] implemented Augmented Reality in a severe game to word learning. However, the game aspect is not balanced with the learning aspect. Therefore, Ferrer stated that balancing the game and learning aspects is necessary [11].

Procedural Content Generation (PCG) was implemented in previous research by Hoosyar [12] and Ar-Rosyid [13] to balance the learning aspect and game aspect. However, in the study conducted by Hoosyar [12], PCG did not include the game aspect. On the other hand, the previous work by Ar-Rosyid [13] cannot adjust the learning content difficulty to player skill.

To adjust the learning content difficulty, Dynamic Difficulty Adjustment (DDA) was

implemented by Jennings and Baldwin [14], [15]. The application of DDA allows the game to adjust the difficulty level automatically. DDA can also keep learners within the Zone of Proximal Development (ZPD) limits. According to Vygotsky, ZPD can make students learn better [16].

The main contribution of the work presented in this paper is summarized as follows: (a) a PCG for balancing the learning aspect and game aspect is proposed; (b) from the proposed PCG, a matrix for reading skills was developed; (c) under the proposed PCG, DDA, and reading skill matrix, a serious game, ABC Adventure was developed, and the game has been tested with the human player via user survey and statistical analysis.

METHODS

In this research, Design Research Methodology (DRM) [17] was adapted as research methodology and illustrated in figure 1.

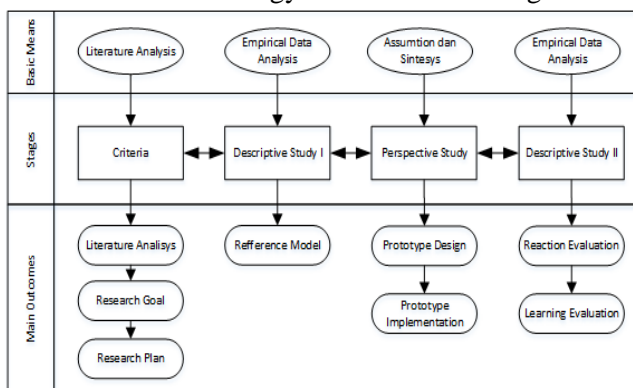


Figure 1. DRM [17]

Based on figure 1, a literature analysis was conducted from the first step. The next step is determining the research goal and setting the research plan. In the second step, a literature review was developed to map the literature study, as seen in figure 6. In the third step, the game prototype was designed and implemented. The game design was adapting the methods used in previous research conducted by Hoosyar [12], Ar-Rosyid [13], and Lach [18]. The game design in this research shown in figure 2.

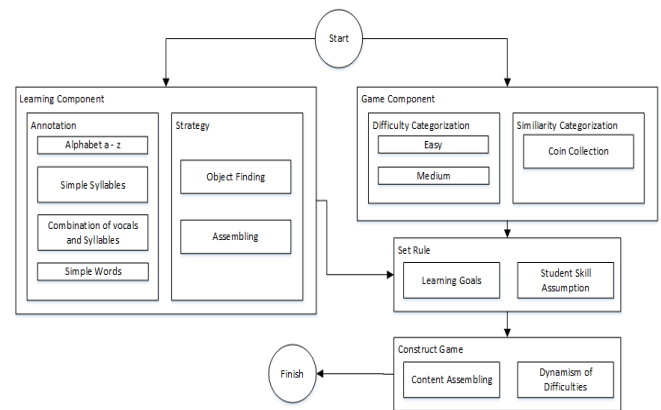


Figure 2. Game Design

After designing the game, the next step was implementing the design into the game. The mobile platform was chosen for the game because the number of mobile gamers is increasing [5], [6]. On the mobile platform, the educational genre is still far below other genres in terms of popularity [7], [8]. Therefore this is a good opportunity to attract the player’s interest in playing learning games. The game was made in Virtual Reality (VR), as shown in figure 3, because based on a survey conducted by the Department of American Education [19], educational game in VR is needed.

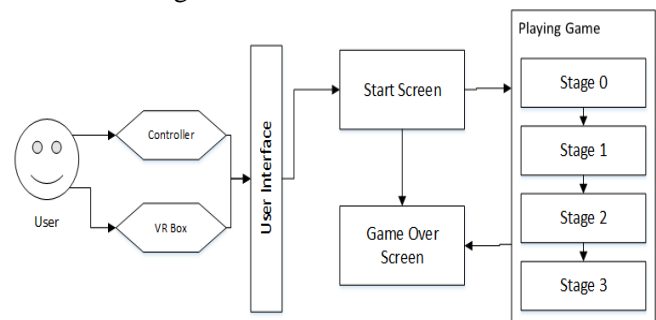


Figure 3. Game Architecture

In the last step, the student’s reaction according to the game playing and its learning outcome were evaluated. The four-level evaluation method proposed by Kirkpatrick [20] as shown in figure 4 below, was adapted for the evaluation method.

This research evaluated the balance between the learning aspect and game aspect by player reaction and learning outcome. Therefore, only two levels of evaluation: Reaction and Learning, were adopted.

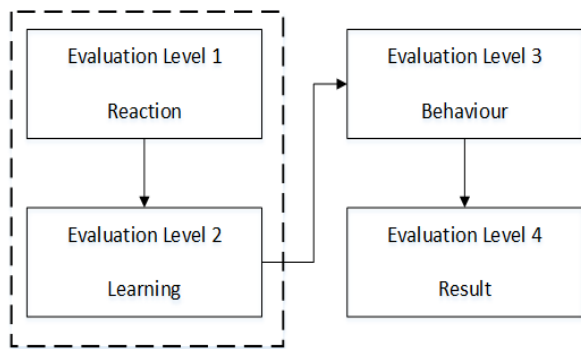


Figure 4. Four Level Evaluation [20]

In this research, game enjoyment was measured as the player’s reaction. To measure player enjoyment EGameFlow method [21] was adopted. From that method, there are eight enjoyment factors: (1) Concentration; the game must provide an activity that encourages the player’s concentration while minimizing stress from learning; (2) Goal Clarity, a task in the game should be clearly explained at the beginning; (3) Feedback, feedback allows the player to determine the gap between the current states of knowledge; (4) Challenge, the game should offer challenges that fit player skills; (5) Autonomy, the learner should enjoy taking the initiative in game playing; (6) Immersion, the game should lead the player into a state of immersion; (7) Social Interaction, a task in the game should become a means for the player to interact socially; (8) Knowledge Improvement, the game should increase the player’s level of knowledge. These eight factors were measured by Likert [22] scale questionnaire from scales 1 to 5, filled with the students.

To measure learning outcomes, the student score from game tasks was used. The student score was measured by t-test [23], and the effectivity of the game was evaluated by calculating the N-Gain Score [24].

A. Balancing Learning Aspect and Game Aspect in Serious Game

A well-designed serious game can connect the learning and game aspects [25]. It has a serious aspect of learning but still enjoyable to play. Therefore, the goals of serious

games are how to connect or balance the two aspects, the learning aspect and the game aspect [26].

A Procedural Content Generation (PCG) Method was adopted to balance the two aspects. PCG is a method or process to generate game content such as terrain, level, map, story, dialogue, task, character, dynamic texture, and weapons based on the specified parameter [27]. There are three common methods in PCG: constructive, simple generate-and-test, and search based, as illustrated in figure 5.

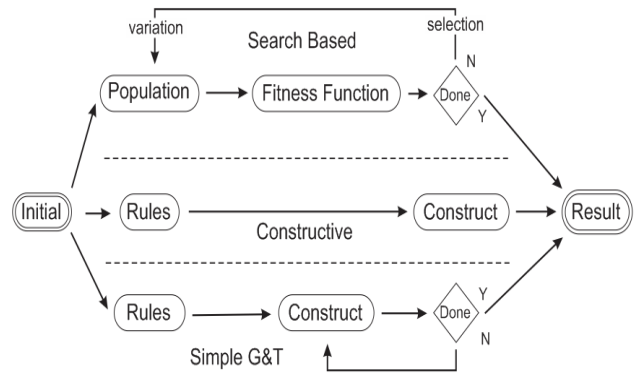


Figure 5. PCG Methods [27]

Based on Hendrikx [28], several common strategies are used to deploy content in a game. The first and most commonly used method is using a Pseudo-Random Number Generator. Furthermore, Lameraz [29] stated a connection between learning attributes and game attributes. One of the learning attributes needed in a serious game is Information Transmission. The Information Transmission attribute is connected to the Challenge Repetition in-game attribute, and the result is remembering. Therefore, to make the student remember the learning materials, Challenge Repetition is necessary for our game.

Hoosyar [12] and Ar-Rosyid [13] conducted research related to balancing the learning aspect and game aspect. The method by Hoosyar generates learning content based on the learning domain, learning goals, and player skill assumption. A genetic algorithm is used to determine good or bad materials. However, Hoosyar [12] did not include the game aspect of

this method. On the other hand, Ar-Rosyid [13] add the game aspect to their PCG method. Based on the method proposed by Ar-Rosyid [13], PCG consists of two parts: the learning aspect and the game aspect. However, game content generated by this method cannot adapt to player skill.

B. Difficulty Adjustment in Serious Game

The two most common ways to automatically adjust the difficulty in the game are using the Genetic Algorithm (GA), and Dynamic Difficulty Adjustment (DDA) approaches. GA offers the potential solution by recombining the data structure [30] and implementing it in varied combinatorial optimization problems [31]. Meanwhile, the DDA approach is a mechanism that will make the game automatically change in a certain way to better suit the players' skill [14] in real-time play [15]. This method aims to give the player better experiments and challenges [32] and maintain the balance in the game [33].

The latest DDA was proposed by Lach [18] by generating game content similar to player characteristics e.g., life, health point, ammo, time, etc. According to this method, the game mechanism will generate game content based on player performance. For example, if the player has a low health point, then the player will get an extra life point if he can complete the task.

C. Reading Skill Methods

The most common method in reading skills is the Glenn Doman method [34]. According to this method, teaching the alphabet is not necessary for developing reading skills. However, reading skills can start with simple words. There are five steps in these methods. The first is to start with simple words. The kids will learn simple words from daily objects or environments in this step. The next step is combining two words. The kids will learn a combination of the simple words from step one. In the third step, the kids will learn a simple sentence from a combination of the previous steps. In the fourth step, the kids will learn the meaning of the sentence, and in the last step, the kids will start to read from the books.

The more systematic method for reading skills is ALBA, proposed by Rasto [35]. In these methods, there are two major steps. The first step is TLT, Tunjuk (pointing), Lafalkan (pronouncing), and Tirukan (imitate). The second step is BMD, Baca (read), Mandiri (autonomous), and Dampingi (accompany). In the TLT step, the kids will learn about the non-capital alphabet from a to z and simple syllables such as ba, bi, bu. After the kids can remember the alphabet and syllables, he can continue to the BMD step. In BMD, the kids will learn simple words such as i-bu or u-da, and then simple words like cu-cu. In the final exercise, the kids will learn sentences

D. Literature Map

A literature map has been developed from the related works discussed before for mapping the related works used in this research. Figure 6 below shows the literature map.

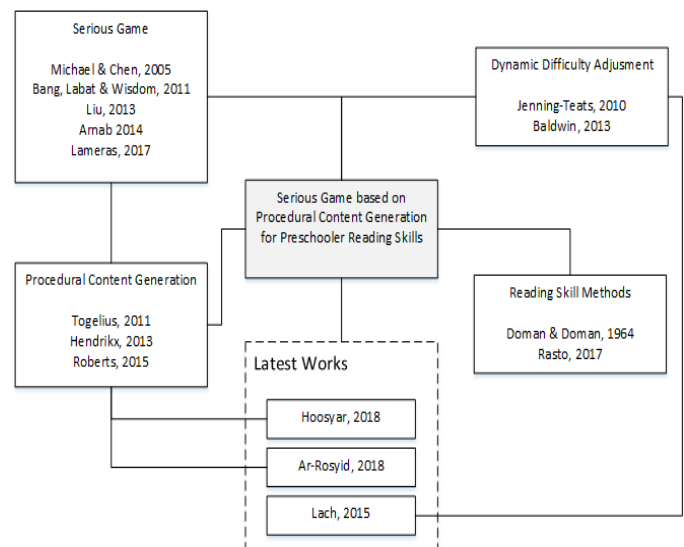


Figure 6. Literature Map

RESULT AND DISCUSSION

From reading skill methods [34], [35] discussed in the previous section, a matrix for reading skill learning, from alphabets to sentence and paragraph reading systematically, shown in Table 1, has been developed.

Table 1. Reading Skill Matrix

Lv	Learning Material	Difficulties
0	Non-capital alphabet a-z	Easy
1	Syllables	Easy
2	Simple words	Easy
3	Combination of letters and Syllables	Easy, Medium, Hard
4	Simple Sentences consist of simple words	Hard
5	Complex Sentences	Hard
6	Paragraph or books	Hard
7	Identifying missing letters from words	Medium
8	Identify object name by images	Hard
9	Capital letters	Easy

The matrix has the 9-level reading skills, and the difficulty was determined based on the curriculum and learning syllabus. In this research, level 0-3 was employed. A constructive method from PCG has been utilized to construct the game content. The rules for constructing the game content are as follows: (1) 26 letters from the alphabet a-z for level 0 were used; (2) For levels 1-3, 5 vocal letters a, i, u, e, o and 18 consonant letters excluding v, x, q was used; (3) Combine the consonant letters from b until z to a, i, u, e, o that is the result are syllables ba, bi, bu, be, bo, until zo; (4) Utilizing the same methods, combine the syllables to generate simple words; (5) Determine the difficulty of the learning materials into easy, medium, or hard based on curriculum and learning syllabus; (6) Coin collection genre was chosen because the genre fits the learning strategy.

Based on the rules, the maximal number of game content generated in this game was identified. Figure 7 illustrates the total generated game content.

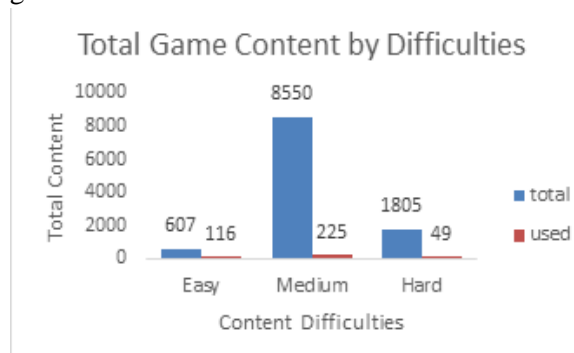


Figure 7. Total Game Content

From figure 7 there is 10.962 game content that can be generated in this serious game, consisting of 607 content easy, 8550 content medium, and 1805 content hard. However, in this research, only 390 game content was used.

There are several rules to deploy the game content and adjust the game’s difficulty. The rules are: (1) Created a content set; consist of 4 content, easy and medium; (2) The first content is easy; (3) The following contents are pseudo-randomly generated, according to easy content can be generated three times at maximal, medium content can be generated two times at maximal; (4) Add life point in the game to add more challenge to the player. The life points start at five if the player completes the task, the life points plus 1 until maximal 6. If the player fails the task, the life points decrease by 1, and the game is over if life points equal 0.

The serious game developed in this research is called ABC Adventure (figure 8). ABC Adventure consists of four stages, from stage 0 to stage 3 (figure 3). In this game, the player will be given letters or syllables as a task. To complete the task, the player has to look for the constituent letters based on the shown task in the game environment.

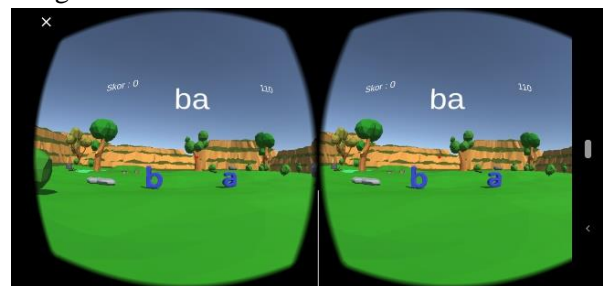


Figure 8. ABC Adventure

In figure 8, it is shown that the task is syllable ba. To complete the task, the player has to look around the letters b and a and choose the correct order. If the player completes the task correctly, they will get a score of 10 per task. This score will be evaluated to measure the learning outcome. A black box test has tested game functionality. The result of the black box test is shown in Table 2.

Table 2. Black Box Test Result

No	Page	Scenario	Result
1	Start Screen	Entering Start Screen	OK
2	Game Over Screen	Entering Game Over Screen	OK
3	Stage 0	Entering Stage 0	OK
4	Stage 1	Entering Stage 1	OK
5	Stage 2	Entering Stage 2	OK
6	Stage 3	Entering Stage 3	OK
7	Content Generation	In-game testing	OK
8	Difficulty Dinamization	In-game testing	OK

An experiment with 32 preschool students has been conducted. Each student plays the game for around 10 minutes at five days. After the students play the game, we interview the student to measure the game's enjoyment. Our game is a single-player game, so the social interaction factor was not included. Several statistical tests were conducted on the questionnaire: validity test [36], reliability test [37], data transform [38], and ANOVA test [39], [40]. From the test result, the autonomy factor was not valid, so it was not included in the next test. A reliability test showed that all questions were reliable. After transforming data, an ANOVA test was conducted, and the result was positive, as shown in Table 3.

Table 3. ANOVA Test Result

No.	Enjoyment Factor	Mean
1	Concentration	3.25
2	Goal Clarity	2.65
3	Feedback	2.44
4	Challenge	3.16
5	Immersion	3.25
6	Knowledge Improvement	2.84

From table 3, On scale 5, 3 factors get a mean score greater than 3.0. The factors were concentration, challenge, and immersion. From this score, it can be concluded that games can make users more concentrated, feel challenged, enjoy the game, and improve knowledge. The mean score was a similar base from between groups sig. value $0.193 > 0.05$.

To evaluate learning outcomes, several statistical tests, including normality test, homogeneity test [41], paired t-test [23], and N-Gain test [24], were conducted. The normality test showed that the data were normally distributed with sig. value 0.186. The homogeneity test also showed that the data were homogenous with sig. value 0.414. From paired t-test, the first experiment's mean score was 47.94, and from the last experiment, the mean score was 82.85, as shown in Table 4.

Table 4. Paired t-Test Result

Paired Samples Statistics				
	Mean	N	Std. Deviation	Std. Error mean
First Experiment	47.94	32	14.90323	2.634
Last Experiment	85.66	32	17.29630	3.057
Paired Samples Correlation				
	N	Correlation	Sig.	
First and Last Experiment	32	0.731	0.000	
Paired Samples Test				
	t	df	Sig. (2-tailed)	
	-16.446	31	0.000	
Pair Differences				
Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference	
			Lower	Upper
-34.909	12.0078	2.12270	-39.238	-30.579

The mean score shows significant differences in learning outcomes after more learning using a serious game. This result also was shown by sig. value $0.000 < 0.005$. Based on the N-Gain test, the serious game was medium effective in improving reading skills. The N-Gain mean score was 0.551 from maximal scale 1 (scores 0.3-0.7 are medium effective). The game did not get good score effectiveness because there were several obstacles while experimenting such as the time range from experiment 1 to the last experiment being too short. To get a good result, the experiment has to conduct for several months. The other thing was the student result in completing the task was not

so different. Many students only scored 10-40 points from experiment 1 to the last.

CONCLUSION

A serious game based on procedural content generation for preschooler reading skills was successfully developed. It gives a new experience to the player by providing varied game content and challenge. According to the player's reaction in the ANOVA result and the player's learning outcome in the t-Test result, the serious game developed in this research can connect and balance the two aspects needed in the educational game: the learning aspect and game aspect.

Several suggestions can be made in future research: (1) Utilizing a database to store PCG-generated game content; (2) Add other game elements such as clues and different terrains; (3) Game developers must concern about game aspects such as instruction, playtime, and game genre to give a better experience in learning and playing; (4) Implementing voice recognition and motion tracking will be helpful in a serious game for reading skills.

REFERENCES

- [1] Pusat Bahasa Departemen Pendidikan Nasional, Indonesian Dictionary. 2008.
- [2] D. Seftiawan, "Masih ada 3,4 Juta Orang Buta Huruf di Indonesia." <https://www.pikiran-rakyat.com/pendidikan/2018/09/04/masih-ada-34-juta-orang-buta-huruf-di-indonesia-429732> (accessed May 25, 2019).
- [3] F. Sidik, "Pengguna Perangkat Mobile di Indonesia Semakin Tinggi ini Datanya." <https://teknologi.bisnis.com/read/20180201/101/733037/pengguna-perangkat-mobile-di-indonesia-semakin-tinggi-ini-datanya> (accessed May 25, 2019).
- [4] F. Novalius, "Indonesia Pengguna Smartphone ke 4 Dunia." <https://economy.okezone.com/read/2018/02/17/320/1860752/indonesia-pengguna-smartphone-ke-4-dunia-begini-tekad-menperin-dongkrak-industri-telematika> (accessed May 25, 2019).
- [5] L. Maulida, "Jumlah Gamer di Indonesia Capai 100 Juta di 2020." <https://www.tek.id/insight/jumlah-gamer-di-indonesia-capai-100-juta-di-2020-b1U7v9c4A> (accessed May 25, 2019).
- [6] Samsung Kidstime, "Mobile Device Usage Among Young Kids A Southeast Asia Study." <https://id.theasianparent.com/hasil-survey-smartphone-yang-mengejutkan> (accessed May 25, 2019).
- [7] Go GLOBE, "Mobile Gaming Industry." <https://www.go-globe.com/blog/mobile-gaming-industry> (accessed May 25, 2019).
- [8] WePC, "Mobile Device Usage Among Young Kids", Accessed: May 25, 2019. [Online]. Available: <https://www.wepc.com/news/video-game-statistics>
- [9] D. Michael and S. Chen, *Serious Games: Games that Educate, Train and Inform*, 1st ed. Thomson Course Technology, 2006.
- [10] D. R. A. Rambli, W. Matcha, and S. Sulaiman, "Fun learning with AR alphabet book for preschool children," in *Procedia Computer Science*, 2013, vol. 25, pp. 211–219. doi: 10.1016/j.procs.2013.11.026.
- [11] J. R. C. Ferrer, "Educational games as stand-alone learning tools and their motivational effect on L2 vocabulary acquisition and perceived learning gains," *British Journal of Educational Technology*, vol. 48, no. 2, pp. 264–278, Dec. 2015.
- [12] D. Hoosyar, M. Yousefi, and H. Lim, "A Procedural Content Generation-Based Framework for Educational Games: Toward a Tailored Data-Driven Game for Developing Early English Reading Skills," *Journal of Educational Computing Research*, vol. 56, no. 2, pp. 293–310, 2018.
- [13] H. Ar Rosyid, M. Palmerlee, and K. Chen, "Deploying Learning Material to Game Content for Serious Education Game Development," *Entertainment Computing*, pp. 1–9, 2018.
- [14] M. J. Teats, G. Smith, and N. W. Fruin, "Polymorph: dynamic difficulty adjustment through level generation," in *PCGames '10: Proceedings of the 2010 Workshop on Procedural Content Generation in Games*, Jun. 2010, pp. 1–4.
- [15] A. Baldwin, D. Johnson, P. Wyeth, and P. Sweetser, "A Framework of Dynamic Difficulty Adjustment in Competitive Multiplayer Video Games," in *Games Innovation Conference (IGIC)*, 2013, pp. 16–19.
- [16] B. du Boulay and R. Luckin, "Modelling Human Teaching Tactics and Strategies for Tutoring Systems," 2001.

- [17] L. T. M. Blessing and A. Chakrabarti, DRM, a design research methodology. Springer London, 2009. doi: 10.1007/978-1-84882-587-1.
- [18] E. Lach, "A Quick Method for Dynamic Difficulty Adjustment of a Computer Player in Computer Games," in *International Conference on Artificial Intelligence and Soft Computing*, 2015, pp. 669–678.
- [19] M. Molnar, "Education Seen as Strong Market for VR and AR By Industry Insiders." <https://marketbrief.edweek.org/marketplace-k-12/education-seen-strong-market-vr-ar-industry-insiders/> (accessed May 25, 2019).
- [20] D. Kirkpatrick, *Evaluating Training Programs: The Four Levels* (3rd Edition), 3rd ed. American Society for Training and Development, 2006.
- [21] F.-L. Fu, R.-C. Su, and S.-C. Yu, "EGameFlow: A scale to measure learners' enjoyment of e-learning games," *Computers & Education*, vol. 52, no. 1, pp. 101–112, Jan. 2009.
- [22] R. Likert, "A technique for the measurement of attitudes," *Archives of Psychology*, 1932.
- [23] Student, "The Probable Error of a Mean," *Biometrika*, pp. 1–25, 1908.
- [24] R. R. Hake, "Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses," *American Journal of Physics*, vol. 66, no. 1, pp. 64–74, 1998.
- [25] B. Huynh-Kim-Bang, J. Wisdom, and J.-M. Labat, "Design Patterns in Serious Games: A Blue Print for Combining Fun and Learning."
- [26] S. Arnab et al., "Mapping learning and game mechanics for serious games analysis," *British Journal of Educational Technology*, vol. 46, no. 2, pp. 391–411, Mar. 2015, doi: 10.1111/bjet.12113.
- [27] J. Togelius, G. N. Yannakakis, K. O. Stanley, and C. Browne, "Search-Based Procedural Content Generation: A Taxonomy and Survey," *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 3, no. 3, pp. 172–186, Sep. 2011, doi: 10.1109/TCIAIG.2011.2148116.
- [28] M. Hendriks, S. Meijer, J. van der Velden, and A. Iosup, "Procedural content generation for games," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 9, no. 1, pp. 1–22, Feb. 2013, doi: 10.1145/2422956.2422957.
- [29] P. Lamas, S. Arnab, I. Dunwell, C. Stewart, S. Clarke, and P. Petridis, "Essential features of serious games design in higher education: Linking learning attributes to game mechanics," *British Journal of Educational Technology*, vol. 48, no. 4, pp. 972–994, Jun. 2017, doi: 10.1111/bjet.12467.
- [30] D. Whitley, "A genetic algorithm tutorial," *Statistics and Computing*, vol. 4, no. 2, Jun. 1994, doi: 10.1007/BF00175354.
- [31] J. E. Beasley and P. C. Chu, "A genetic algorithm for the set covering problem," *European Journal of Operational Research*, vol. 94, no. 2, pp. 392–404, Oct. 1996, doi: 10.1016/0377-2217(95)00159-X.
- [32] R. Hunicke, "The case for dynamic difficulty adjustment in games," in *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology - ACE '05*, 2005, pp. 429–433. doi: 10.1145/1178477.1178573.
- [33] E. Adam, "Balancing Games With Positive Feedback," 2002.
- [34] G. Doman and J. Doman, *How to Teach Your Baby to Read*. Gentle Revolution Press, 1964.
- [35] Rasto, *Mudah Belajar Membaca Permulaan Teknik Alba*. Rizqi Press, 2017.
- [36] K. Pearson, "VII. Note on regression and inheritance in the case of two parents," *Proceedings of the Royal Society of London*, vol. 58, no. 347–352, pp. 240–242, Dec. 1895, doi: 10.1098/rspl.1895.0041.
- [37] L. J. Cronbach, "Coefficient alpha and the internal structure of tests," *Psychometrika*, vol. 16, no. 3, pp. 297–334, Sep. 1951, doi: 10.1007/BF02310555.
- [38] Sedarmayanti and S. Hidayat, *Metodologi Penelitian*. Bandung: Mandar Maju, 2002.
- [39] R. A. Fisher, "XV.—The Correlation between Relatives on the Supposition of Mendelian Inheritance.," *Transactions of the Royal Society of Edinburgh*, vol. 52, no. 2, pp. 399–433, Jul. 1919, doi: 10.1017/S0080456800012163.
- [40] Sugiyono, *Metode Penelitian Kualitatif dan R&D*. Bandung: Alfabeta, 2007.
- [41] P. A. Games, H. B. Winkler, and D. A. Probert, "Robust Tests for Homogeneity of Variance," *Educational and Psychological Measurement*, vol. 32, no. 4, pp. 887–909, Dec. 1972, doi: 10.1177/001316447203200404.