



Development of the Blood Race Educational Board Game as a Science Learning Medium on the Human Circulatory System

Della Afreadita Kusuma^a, Fathonah^b, Lara Puspita^c, Septi Olivia^d,
Tri Nurleli^e, Tiwi Marlita^f, Meirita Sari^g

^{abcdefg}Universitas Islam Negeri Fatmawati Sukarno Bengkulu, Indonesia

Abstract

Learning the human circulatory system is often difficult for students because the process cannot be observed directly. This study aimed to develop Blood Race, an educational board game for science learning, and to test its validity, practicality, and effectiveness. This quantitative development research followed the ADDIE model. Validity was assessed by three content experts and three media experts using Aiken's V, practicality was measured through a response questionnaire completed by 32 eighth-grade students, and effectiveness was tested using a one-group pretest-posttest design. Aiken's V was 0.83 for content and 0.87 for media, with feasibility percentages of 86.67% and 89.44%. Student responses reached 84.52%, categorized as very practical. The mean score increased from 45.47 to 82.81 ($t(31) = 12.61$; $p < 0.001$; $d = 2.73$), with a mean N-Gain of 0.68 (medium category) and classical mastery of 81.25%. Therefore, Blood Race is valid, practical, and effective in improving science learning outcomes on the human circulatory system; however, the absence of a comparison group means that the findings are preliminary evidence.

Keywords: educational board game; Blood Race; science learning media; human circulatory system; N-Gain.

INTRODUCTION

Natural science (*Ilmu Pengetahuan Alam*, IPA) requires students not only to memorize facts but also to understand processes and the relationships among concepts. The human circulatory system is a challenging topic because most of its processes take place inside the body and cannot be observed directly. Students must imagine the path of blood through the heart, lungs, and body, distinguish pulmonary from systemic circulation, and relate blood components to the needs of the body. Izza, Sukamti, and Winahyu (2021) reported that students still hold misconceptions about this topic. Instruction dominated by lectures and static pictures tends to leave students passive, even though instructional media are meant to clarify messages, overcome limited experience, and stimulate motivation to learn (Arsyad, 2017).

These conditions call for media that give students an opportunity to trace the path of blood step by step, test their understanding, and receive immediate feedback. The learning outcomes for science in Phase D emphasize the ability to explain the relationship between the structure and function of organs, so active and meaningful learning activities are highly relevant. In practice, teachers often face limited class time, large classes, and uneven access to technology. Media that are simple, durable, and playable in small groups are therefore a realistic choice.

Educational games are one alternative medium that places students as active participants in learning. Plass, Homer, and Kinzer (2015) explain that game-based learning can be examined from cognitive, motivational, affective, and sociocultural perspectives. The meta-analysis by Clark, Tanner-Smith, and Killingsworth (2016) showed that digital games can enhance learning compared with conditions without games, while the meta-analysis by Sailer and Homner (2020) found that gamification has positive effects on cognitive learning outcomes, motivation, and behavior, with small to medium effect sizes. Hamari et al. (2016) showed that challenge in games is related to students' engagement and immersion, which in turn are related to learning outcomes. For board games specifically, the systematic review by Noda, Shiotsuki, and Nakao (2019) of 27 studies concluded that board games can be expected to improve understanding of knowledge, interaction among participants, and motivation.

Media development studies on the circulatory system have been carried out in various forms. Fajarianti (2023) developed a Wordwall-based educational game for fifth-grade elementary school students with an N-Gain of 0.88. Munawaroh (2024) developed the role-play-based game *Adventure of Blood* with 23 fifth-grade students and reported an increase in the mean score from 66.96 to 82.83. Sangi and Bata (2023) developed augmented reality-based media, whereas Hardiyanti, Mustami, and Mu'nisa (2020) developed a puzzle game based on Construct 2 for eleventh-grade senior high school students. For another organ system, Hapsari and Mintohari (2025) developed the *Truth or Dare Board Game* based on *Air Journey* for the respiratory system and obtained a mean N-Gain of 0.67.

Double circulation has a distinctive feature, namely two consecutive pathways. Blood is pumped from the right ventricle to the lungs and returns to the left atrium (pulmonary circulation), and is then pumped from the left ventricle to the whole body and returns to the right atrium (systemic circulation). This sequential character makes the topic well suited to a board game track, because the position of a pawn can represent the position of blood within the circulatory system.

A review of these studies reveals two gaps. First, most of the products on the circulatory system reviewed by the authors are digital and were tested in elementary or senior high schools, so physical board games for junior high school students are still rarely reported. Board games, in contrast, need no devices or internet connection and encourage face-to-face cooperation in groups. Second, reports of effectiveness generally stop at the difference between means or a single N-Gain value, whereas checks of statistical assumptions, effect sizes, the distribution of individual N-Gain scores, and gains

on each content indicator are seldom presented together. This statement is based on the authors' literature search and does not rule out the existence of other studies that have not been reached.

This study addresses these gaps through Blood Race, an educational board game whose track is designed to resemble double circulation. Students move pawns shaped like blood components from the heart to the lungs, back to the heart, and then to the whole body, while answering quiz cards and resolving event cards that cover blood components, circulatory organs, blood types, and circulatory disorders. In this way, the path of blood becomes the mechanism of the game itself rather than merely a visual theme. A further novelty is the comprehensive quantitative evaluation: validity is tested with Aiken's *V*, practicality with a response questionnaire, and effectiveness with a paired-samples *t*-test, effect size, and N-Gain broken down to the indicator level. This study hypothesizes that learning with Blood Race significantly increases students' learning outcome scores and yields an N-Gain of at least the medium category.

This study aims to (1) describe the development of Blood Race, (2) test its validity according to content experts and media experts, (3) test its practicality based on students' responses, and (4) test its effectiveness in improving science learning outcomes through analysis of pretest, posttest, and N-Gain data. Theoretically, the results enrich the evidence on board games as science learning media; practically, the product offers teachers an active, collaborative medium that does not depend on electronic devices.

RESEARCH METHOD

A. Research Type and Design

This study is research and development (R&D) with a quantitative approach (Sugiyono, 2019). The development procedure followed the ADDIE model, which consists of five stages, namely analysis, design, development, implementation, and evaluation (Branch, 2009). The effectiveness of the product was tested using a pre-experimental one-group pretest-posttest design with the scheme $O_1 - X - O_2$ (O_1 = pretest, X = learning with Blood Race, O_2 = posttest). This design was chosen because the study focused on an initial test of the feasibility of the product in one intact class without a comparison class.

B. Research Subjects and Location

The field trial was conducted at MTs Pancasila in Bengkulu City in the odd semester of the 2026/2027 academic year. The subjects were 32 students in one eighth-grade class, selected purposively based on the readiness of the learning schedule and the approval of the science teacher. The test instrument was tried out with 30 ninth-grade students at the same school who had already studied the circulatory system. Product validation was carried out by three content experts (two biology education lecturers and one experienced science teacher) and three media experts (two educational technology lecturers and one educational game developer). All participants gave their consent, and the data are reported without personal identities.

C. Development Procedure

The development stages and their outputs are presented in Table 1. The field trial was carried out in three meetings of 2 × 40 minutes each. The first meeting was used for the pretest and an orientation to the rules of the game. The second and third meetings were used to play Blood Race in eight groups of four students, with the teacher as facilitator. The posttest and the response questionnaire were administered at the end of the third meeting.

Each validator played one round of the game, examined the board and the card sets, and filled in the validation sheet together with suggestions for improvement. At the implementation stage, the teacher explained the rules, managed playing time, and mediated disagreements over answers without changing the content of the cards. The pretest was completed individually in 30 minutes without discussing the answers so that the effect of test repetition would be limited.

Table 1. Stages of development in the ADDIE model and their outputs

Stage	Activities	Output
<i>Analysis</i>	Classroom observation, teacher interviews, and analysis of learning outcomes and student characteristics	Media needs and learning objectives
<i>Design</i>	Designing the game flow and rules, board layout, cards, and the test and questionnaire blueprints	Board design, rules, and blueprints
<i>Development</i>	Prototype production, expert validation, revision, and instrument trial	Validated prototype and final instruments
<i>Implementation</i>	Three-meeting field trial: pretest, gameplay, posttest, and response questionnaire	Learning outcome and response data
<i>Evaluation</i>	Quantitative data analysis, feasibility conclusions, and final revision	Feasibility conclusions and final product

Source: Adapted from Branch (2009)

D. Research Instruments

This study used three types of instruments. First, the content expert and media expert validation sheets each consisted of 12 items rated on a scale of 1 to 5. Second, the learning outcome test was a 20-item multiple-choice test with four options based on the five indicators in Table 2. Each correct answer was scored 5, giving a maximum score of 100, and the individual mastery criterion was set at 75 according to the school's criterion. The same test was used for the pretest and the posttest. Third, the student response questionnaire consisted of 16 items covering four aspects, namely interest, ease of use, benefit for understanding the material, and cooperation and learning motivation, with a Likert scale of 1 to 4 (Sugiyono, 2019).

A total of 25 test items were originally written, five for each indicator, and their content was validated by two content experts. From the trial results, valid items with an appropriate difficulty level were selected, four for each indicator, which gave the 20 final items. The questionnaire was content-validated by the same experts, and its internal consistency was calculated using Cronbach's alpha based on the field trial data.

Table 2. Blueprint of the circulatory system learning outcome test

Indicator	Cognitive level (number of items)	Item numbers	Total
Explaining blood components and their functions	C1 (2), C2 (2)	1–4	4
Explaining the structure and function of circulatory organs	C1 (1), C2 (3)	5–8	4
Analyzing the pathways of systemic and pulmonary circulation	C2 (1), C3 (1), C4 (2)	9–12	4
Applying the concepts of blood types and transfusion	C3 (4)	13–16	4
Analyzing circulatory disorders and efforts to maintain health	C3 (2), C4 (2)	17–20	4
Total	C1 (3), C2 (6), C3 (7), C4 (4)	1–20	20

Source: Researchers' processed data (2026)

E. Data Analysis Technique

Product validity was analyzed using Aiken's V coefficient (Aiken, 1985) with the formula $V = \sum s / [n(c - 1)]$, where s is the rating minus the lowest rating score, n is the number of raters, and c is the number of rating categories. The percentages of feasibility and practicality were calculated with $P = (\text{total obtained score} / \text{total maximum score}) \times 100\%$. Item quality was analyzed through the corrected item-total correlation, compared with the r table value of 0.361 ($N = 30$; $\alpha = 0.05$), the difficulty index, and Kuder-Richardson 20 (KR-20) reliability (Arikunto, 2013).

Learning outcome data were analyzed descriptively and then tested for normality using the Shapiro-Wilk test (Shapiro & Wilk, 1965) on the pretest, posttest, and difference scores. The null hypothesis (H_0) was that there is no difference between the mean pretest and posttest scores. The test was a two-tailed paired-samples t -test at a significance level of 0.05, complemented by a 95% confidence interval and the effect size Cohen's d calculated with the pooled standard deviation (Cohen, 1988). The Wilcoxon signed-rank test was used as a sensitivity test. The improvement in learning outcomes was calculated with normalized gain (Hake, 1998) using $g = (\text{posttest score} - \text{pretest score}) / (\text{maximum score} - \text{pretest score})$. The interpretation criteria are presented in Table 3.

A product was declared feasible if it met four criteria. First, the mean Aiken's V was at least 0.80, with a feasibility percentage in the feasible category or higher. Second, the practicality percentage was in the practical category or higher. Third, the paired-samples t -test showed a significant improvement with a mean N-Gain of at least the medium category. Fourth, classical mastery was at least 75%, meaning that at least 75% of students obtained a score of 75 or higher.

Table 3. Criteria for interpreting research data

Parameter	Range	Category
Aiken's V coefficient	$V \geq 0.80$	Valid
	$0.60 \leq V < 0.80$	Fairly valid

Parameter	Range	Category
	$V < 0.60$	Not valid
Feasibility and practicality percentage	81–100%	Very feasible/practical
	61–80%	Feasible/practical
	41–60%	Fairly feasible/practical
	21–40%	Less feasible/practical
	$\leq 20\%$	Not feasible/practical
N-Gain (g)	$g \geq 0.70$	High
	$0.30 \leq g < 0.70$	Medium
	$g < 0.30$	Low
Effect size (d)	$d \geq 0.80$	Large
	$0.50 \leq d < 0.80$	Medium
	$0.20 \leq d < 0.50$	Small
Difficulty index (DI)	$DI > 0.70$	Easy
	$0.30 \leq DI \leq 0.70$	Medium
	$DI < 0.30$	Difficult

Source: Aiken (1985), Cohen (1988), Hake (1998), and Arikunto (2013)

RESULTS AND DISCUSSION

A. Product Development Results

At the analysis stage, classroom observations and interviews with the science teacher showed that instruction on the circulatory system still relied on lectures, textbook pictures, and practice questions. Students tended to memorize the names of organs without understanding the order of blood flow, particularly the difference between pulmonary and systemic circulation. An analysis of the Phase D learning outcomes confirmed the need for activities that connect the structure and function of organs with the processes that take place in them. It was therefore determined that the medium must (1) visualize the flow of blood in sequence, (2) involve all members of a group, (3) provide direct feedback, and (4) be easy to make and use without electronic devices.

The design and development stages produced Blood Race, a board game for three to four players in one group. The 50 × 50 cm board shows the heart at the center as the start and finish point, an artery path on the upper side, a vein path on the lower side, and two exchange zones, namely the lungs for pulmonary circulation and the whole body for systemic circulation (Figure 1). Each player moves a pawn shaped like a blood component (erythrocyte, leukocyte, platelet, or plasma) according to the number rolled on the die. The squares on the track have five functions: quiz squares require players to answer a question card, action squares trigger a physiological event such as a narrowed blood vessel that makes the player move back two steps, fact squares give brief information, obstacle squares represent circulatory disorders, and vein points serve as temporary resting places.



Source: Developer documentation (2026)

Figure 1. Design of the Blood Race game board (the labels on the board are in Indonesian, the language used in the field trial)

The game set consists of one board, four pawns, one die, 40 quiz cards, 16 action cards, 12 fact cards, 8 obstacle cards, score tokens, and a teacher's guidebook. The quiz cards are grouped according to the five indicators in Table 2 so that every game covers all subtopics. A player who answers a quiz card correctly earns two score tokens, whereas a player who returns to the heart after passing through the lung zone and the body zone earns a bonus of five tokens. The game ends when one player completes one cycle of double circulation or when the 25-minute time limit runs out, and the winner is the player with the most score tokens. These rules deliberately combine a race with mastery of concepts, so that speed alone is not enough to win without understanding the material.

Validator feedback included simplifying the wording of five quiz cards, adding illustrations to the action cards, and emphasizing the symbols that show the direction of blood flow on the board. All of this feedback was incorporated before the field trial.

At the implementation stage, eight groups played Blood Race under the teacher's facilitation over two meetings. Each group completed one to two rounds in each meeting, and the teacher explained the rules, managed turns, and reinforced answers on quiz cards that were still incorrect. All group members were involved because every turn required the player to read the card, answer, and explain the reasoning to the group.

B. Product Validity

The validation results are presented in Table 4. The mean Aiken's V was 0.83 for the content experts and 0.87 for the media experts, both in the valid category because they exceeded the threshold of 0.80. The feasibility percentages reached 86.67% and 89.44%, respectively, placing both in the very feasible category. On the content aspect, alignment with the learning outcomes obtained the highest value ($V = 0.88$), whereas the quality of

questions and feedback obtained the lowest ($V = 0.79$), slightly below the threshold of 0.80, with the lowest item scoring 0.67. This is in line with the validators' feedback about quiz card wording that was not yet concise, which was corrected before implementation. On the media aspect, physical quality and durability obtained the highest value ($V = 0.90$), followed by visual appearance and by ease of use and material safety, both at 0.85.

Table 4. Results of content expert and media expert validation

Aspect	Items	V	V category	P (%)	Feasibility
Content experts					
Alignment with learning outcomes	4	0.88	Valid	90.00	Very feasible
Concept accuracy	4	0.83	Valid	86.67	Very feasible
Quality of questions and feedback	4	0.79	Fairly valid	83.33	Very feasible
Mean, content experts	12	0.83	Valid	86.67	Very feasible
Media experts					
Visual appearance of the board and components	4	0.85	Valid	88.33	Very feasible
Ease of use and material safety	4	0.85	Valid	88.33	Very feasible
Physical quality and durability	4	0.90	Valid	91.67	Very feasible
Mean, media experts	12	0.87	Valid	89.44	Very feasible

Source: Research data (2026)

This level of validity is similar to that of products in previous studies. Fajarianti (2023) reported media validation of 83.5% and content validation of 84.4% for the Wordwall game, Munawaroh (2024) obtained 91% from content experts and 89% from media experts for *Adventure of Blood*, and Hapsari and Mintohari (2025) obtained 93% and 90% for the *Truth or Dare Board Game*. The feasibility percentages of Blood Race, 86.67% and 89.44%, are thus comparable. Aiken's V coefficients calculated from three raters per expert group have limited statistical power, so these results are better read as evidence of initial feasibility of the content and appearance than as final evidence of product quality.

The high V value on the concept accuracy aspect is important because conceptual errors in a game medium can reinforce existing misconceptions (Izza et al., 2021).

C. Quality of the Learning Outcome Test Instrument

The results of the instrument trial with 30 students are presented in Table 5. Of the 25 items tried out, 21 were declared valid because their corrected item-total correlations exceeded the r table value of 0.361, while 4 items were dropped. The 20 items with the highest correlations and a balanced representation of the indicators were chosen as the final items, four for each indicator; one other valid item was not used because its indicator was already represented by four items with higher discrimination. The mean difficulty index of the final items was 0.59 (medium category), with 15 items in the medium category, 5 in the easy category, and none in the difficult category. The KR-20 reliability coefficient of 0.88 indicates high internal consistency.

Table 5. Summary of the test instrument trial results

Parameter	Result
Number of items tried out	25
Number of valid items ($r \geq 0.361$)	21
Number of invalid items	4
Number of final items (four items per indicator)	20
Mean corrected item-total correlation of final items	0.48
Mean difficulty index of final items	0.59 (medium)
Difficulty distribution: easy / medium / difficult	5 / 15 / 0
KR-20 reliability coefficient	0.88 (high)

Source: Research data (2026)

D. Effectiveness of Blood Race on Learning Outcomes

1. Descriptive statistics of pretest and posttest scores

Learning outcome data from 32 students are presented in Table 6. The mean pretest score was 45.47 ($SD = 15.21$), while the mean posttest score was 82.81 ($SD = 11.98$), a mean difference of 37.34 points. All students improved their scores, with the smallest improvement being 10 points and the largest 70 points, so no student's score decreased or stayed the same. The minimum score rose from 15 to 60, and the maximum score rose from 75 to 100.

Table 6. Descriptive statistics of pretest and posttest scores

Statistic	Pretest	Posttest	Difference
Number of students (N)	32	32	32
Mean (M)	45.47	82.81	37.34
Standard deviation (SD)	15.21	11.98	16.75
Minimum score	15.00	60.00	10.00
Maximum score	75.00	100.00	70.00
Median	45.00	82.50	35.00
Mastered (score ≥ 75)	1 (3.13%)	26 (81.25%)	–
Not yet mastered (score < 75)	31 (96.88%)	6 (18.75%)	–

Source: Research data (2026)

On the pretest, only 1 student (3.13%) reached the mastery criterion, whereas on the posttest 26 students (81.25%) did, so the classical mastery criterion of at least 75% was met. The smaller standard deviation of the posttest shows that the gap between students narrowed.

2. Normality test and hypothesis test

Data normality was examined with the Shapiro-Wilk test (Table 7). The p values for the pretest, posttest, and difference scores were 0.513, 0.084, and 0.424, respectively. All are greater than 0.05, so the data are considered normally distributed and the requirement for the paired-samples t -test, namely normality of the differences, is met. The t -test

showed $t(31) = 12.61$ with $p < 0.001$, which exceeds the two-tailed critical value of 2.040 at $\alpha = 0.05$, so H_0 is rejected.

Table 7. Results of the normality test and hypothesis test

Test	Variable	Statistic	p	Decision
Shapiro-Wilk	Pretest	$W = 0.970$	0.513	Normal
Shapiro-Wilk	Posttest	$W = 0.942$	0.084	Normal
Shapiro-Wilk	Difference	$W = 0.967$	0.424	Normal
Paired-samples t -test	Posttest – pretest	$t(31) = 12.61$	< 0.001	H_0 rejected
Wilcoxon	Posttest – pretest	$W = 0$	< 0.001	H_0 rejected

Source: Research data (2026)

The mean difference of 37.34 points (standard error = 2.96) had a 95% confidence interval of [31.30, 43.38], so even the lower bound is far above zero. Cohen's d of 2.73 ($d_z = 2.23$) falls in the large category. The Wilcoxon test, which does not require normality, gave the same conclusion ($p < 0.001$), so the finding does not depend on distributional assumptions. There is therefore a significant difference between learning outcomes before and after learning with Blood Race, and the research hypothesis is supported.

3. N-Gain analysis

The mean individual N-Gain was 0.68 ($SD = 0.23$; range 0.22–1.00), which is in the medium category and only 0.02 below the boundary of the high category. As shown in Table 8, 17 students (53.13%) were in the high category, 13 (40.63%) in the medium category, and 2 (6.25%) in the low category, so 93.75% of the students obtained a medium or high gain.

Table 8. Distribution of individual N-Gain categories

Category	Range	Number	Percentage (%)
High	$g \geq 0.70$	17	53.13
Medium	$0.30 \leq g < 0.70$	13	40.63
Low	$g < 0.30$	2	6.25
Total		32	100.00

Source: Research data (2026)

An additional analysis showed that pretest scores were strongly and negatively correlated with difference scores ($r = -0.72$), meaning that students with low prior knowledge gained more points. However, the correlation between pretest scores and N-Gain was close to zero ($r = -0.08$), which indicates that the improvement relative to the room available for growth did not depend on initial ability. Blood Race therefore appears to be proportionally beneficial for students with both low and high initial ability. The two students in the low category need special guidance or remedial teaching.

N-Gain complements the t -test because the t -test only shows whether the mean difference differs from zero, whereas N-Gain expresses the improvement relative to the maximum that each student could achieve (Hake, 1998). A student with a pretest score

of 75, for example, can improve by only 25 points, so an improvement of 20 points still yields an N-Gain of 0.80 (high category).

4. Improvement on each indicator

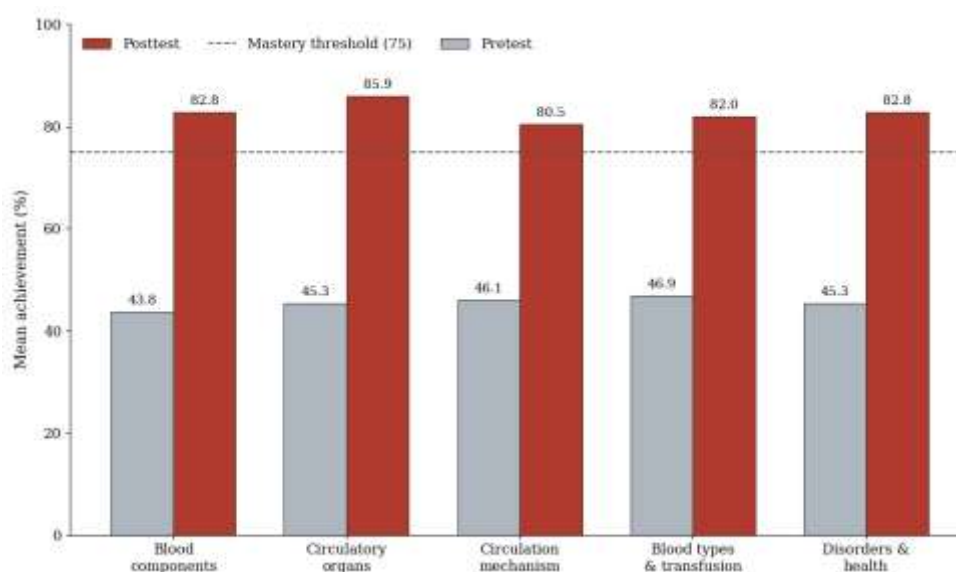
Table 9 and Figure 2 present the achievement on each indicator. The mean posttest achievement on all indicators ranged from 80.47% to 85.94%, so all of them exceeded the mastery threshold of 75. The highest N-Gain was obtained on the circulatory organs indicator ($g = 0.74$; high category), while the lowest was obtained on the mechanism of systemic and pulmonary circulation indicator ($g = 0.64$; medium category).

Pretest scores were relatively uniform across indicators (43.75% to 46.88%), so differences in N-Gain among indicators cannot be explained by large differences in initial ability and more likely reflect how well the playing experience supported each indicator. The largest gain was on the circulatory organs indicator (40.63 percentage points) and the smallest on the mechanism of systemic and pulmonary circulation indicator (34.38 percentage points). Each indicator is represented by only four items, so estimates per indicator are less precise than the total score.

Table 9. Achievement and N-Gain on each indicator

Indicator	Pretest (%)	Posttest (%)	Difference	g	Category
Blood components and their functions	43.75	82.81	39.06	0.69	Medium
Circulatory organs	45.31	85.94	40.63	0.74	High
Mechanism of systemic and pulmonary circulation	46.09	80.47	34.38	0.64	Medium
Blood types and transfusion	46.88	82.03	35.16	0.66	Medium
Disorders and efforts to maintain health	45.31	82.81	37.50	0.69	Medium

Source: Research data (2026)



Source: Research data (2026)

Figure 2. Comparison of mean pretest and posttest achievement on each indicator

E. Practicality of Blood Race

The results of the student response questionnaire are presented in Table 10. The overall practicality percentage reached 84.52%, in the very practical category. All aspects were very practical, with the highest score on interest in the game (86.91%) and the lowest on benefit for understanding the material (83.01%). The questionnaire had high internal consistency (Cronbach's alpha = 0.81). This pattern suggests that students enjoy the game more readily than they recognize its learning benefits, so teachers should close each session with a brief reflection that connects the playing experience with the concepts being learned.

Table 10. Results of the student response questionnaire

Aspect	Obtained score	Maximum score	P (%)	Category
Interest in the game	445	512	86.91	Very practical
Ease of use and rules	435	512	84.96	Very practical
Benefit for understanding the material	425	512	83.01	Very practical
Cooperation and learning motivation	426	512	83.20	Very practical
Overall	1,731	2,048	84.52	Very practical

Source: Research data (2026)

This level of practicality is comparable to that of previous studies: Fajarianti (2023) reported 88.6%, Munawaroh (2024) reported 95%, and Hapsari and MintoHari (2025) reported 91%. The percentage for Blood Race is slightly lower but falls in the same category. A possible cause is that the rules include several types of squares, so students need one initial round to understand them, although the ease of use and rules aspect still reached 84.96%.

F. Summary of Product Feasibility

Table 11 summarizes the four feasibility criteria set in the research method section, namely validity, practicality, improvement in learning outcomes, and classical mastery. All criteria were met, so Blood Race is declared feasible as a science learning medium on the human circulatory system. The N-Gain only met the lower limit of the medium category, so there is still room for improvement, particularly on the mechanism of systemic and pulmonary circulation indicator.

Table 11. Summary of the fulfillment of product feasibility criteria

Criterion	Result	Decision
Content expert validity (<i>V</i>)	0.83 (86.67%)	Met
Media expert validity (<i>V</i>)	0.87 (89.44%)	Met
Practicality (response questionnaire)	84.52% (very practical)	Met
Improvement in learning outcomes (<i>t</i> -test)	$t(31) = 12.61; p < 0.001$	Met

Criterion	Result	Decision
N-Gain at least in the medium category	0.68 (medium)	Met
Classical mastery $\geq 75\%$	81.25%	Met

Source: Research data (2026)

G. Discussion

1. Factors underlying the effectiveness of Blood Race

The increase in the mean score of 37.34 points, the effect size of $d = 2.73$, and the mean N-Gain of 0.68 are consistent with the hypothesis that learning with Blood Race significantly improves learning outcomes. The most likely explanation lies in the design of the product. A board track that imitates double circulation turns a previously abstract flow into a sequence of positions that can be seen, touched, and traversed by a pawn. Students no longer merely memorize that blood flows from the right ventricle to the lungs and back to the left atrium, but experience it as a sequence of steps on the board. The highest N-Gain on the circulatory organs indicator (0.74) is consistent with this explanation, because the heart and blood vessels are the most prominent spatial markers on the board. This is in line with the function of learning media to clarify messages and overcome the limits of direct experience (Arsyad, 2017).

The second factor is the quiz cards. Each time a pawn stops on a quiz square, the player must recall and apply a concept and then receives feedback from the facilitator or group mates. Because the cards are distributed evenly across the five indicators, practice opportunities are spread over all subtopics, which is consistent with the relatively balanced N-Gain across indicators (0.64 to 0.74). However, this explanation is conjectural because the study did not separately measure the frequency and quality of feedback.

The third factor is the motivational and social aspect of the game. Plass et al. (2015) emphasize that educational games work not only through cognitive but also through motivational, affective, and sociocultural perspectives. In Blood Race, the obstacle squares and score tokens provide challenge, while playing in small groups creates interaction and discussion of answers. Hamari et al. (2016) link challenge in games to engagement and immersion, and the cooperation and learning motivation aspect of the questionnaire reached 83.20% (very practical). This is in line with the review by Noda et al. (2019), which concluded that board games can improve understanding of knowledge, interaction, and motivation.

The mechanism of systemic and pulmonary circulation indicator, in contrast, had the lowest N-Gain (0.64) even though it is the core of the track design. This indicator contains items at cognitive levels C2 to C4 that require reasoning about sequence, direction, and differences in blood content, and therefore needs more repetition than indicators that involve recognition. With 25 minutes per round, students probably did not encounter quiz cards on this indicator often enough. Adding C3 and C4 quiz cards for this indicator and closing the session with an activity in which students compose a map of the blood flow could improve this.

The fairly wide spread of individual N-Gain scores ($SD = 0.23$) shows that the benefit of the game is not uniform. The two students with low N-Gain may have been less active in their groups, had difficulty reading the cards, or had few turns on quiz squares, but this cannot be verified because individual activity was not systematically observed.

2. Consistency with and differences from previous studies

The results of this study are in line with previous studies on similar topics. Munawaroh (2024) reported an increase in the mean score from 66.96 to 82.83 (a difference of 15.87) with a significant paired-samples *t*-test, whereas this study shows an increase from 45.47 to 82.81. The larger difference here is largely related to the lower initial score, which leaves more room for improvement, in addition to differences in grade level, instruments, and length of treatment. The N-Gain of 0.68 is close to the 0.67 reported by Hapsari and Mintohari (2025) for a board game on the respiratory system, but lower than the 0.88 reported by Fajarianti (2023) for the Wordwall game; direct comparisons should be made with caution.

The effect size in this study (2.73) is much larger than the small to medium effect sizes reported in the meta-analyses of digital games by Clark et al. (2016) and of gamification by Sailer and Homner (2020). This should not be interpreted to mean that Blood Race is stronger than the interventions analyzed in those meta-analyses. Meta-analyses compare treatment groups with comparison groups, whereas the one-group pretest-posttest design has no comparison group, so the score difference also includes the influence of other learning during the study period, maturation, and test repetition. The results are therefore better interpreted as preliminary evidence of effectiveness rather than causal evidence.

3. Implications and limitations

Practically, teachers can use Blood Race after introducing the concepts as a structured practice activity, with heterogeneous groups, about 25 minutes per round, and a reflection session at the end. The low production cost and the absence of any need for electricity or devices make the product usable in schools with limited facilities.

Theoretically, this study shows that making the structure of a concept into the mechanism of a game, and not merely a visual theme, is a promising design strategy for sequential topics such as blood circulation. It also shows the value of fuller reporting (assumption checks, effect sizes, individual N-Gain, and analysis per indicator), which reveals, for instance, that the indicator at the core of the game design does not necessarily produce the highest improvement.

This study has several limitations. First, the one-group pretest-posttest design does not control threats to internal validity such as other experiences during the study, maturation, and test repetition, because the pretest and posttest used the same set of questions. Second, the trial was conducted in one class in one school over three meetings, so generalization and long-term durability of learning outcomes cannot be ascertained. Third, the novelty effect of the game may raise motivation only temporarily. Fourth, learning outcomes were measured only in the cognitive domain, and expert validation involved a limited number of raters. Further research is recommended to use a quasi-

experimental design with a control class, parallel test forms, retention measurement, and subjects from several schools.

CONCLUSION

This development study produced Blood Race, an educational board game with a track that imitates double circulation, equipped with quiz, action, fact, and obstacle cards, for learning the human circulatory system in science. The product was declared valid by content experts and media experts, with Aiken's V of 0.83 and 0.87, respectively, and feasibility percentages of 86.67% and 89.44% (very feasible). It is also practical based on the responses of 32 students, with an overall percentage of 84.52% (very practical).

The product is effective in improving learning outcomes. The mean score increased from 45.47 on the pretest to 82.81 on the posttest, and the paired-samples t -test showed a significant difference ($t(31) = 12.61$; $p < 0.001$) with a large effect size ($d = 2.73$). The mean N-Gain of 0.68 is in the medium category, with 93.75% of students in the medium or high category, and classical mastery rose from 3.13% to 81.25%. Improvement occurred on all five content indicators, with the highest N-Gain on the circulatory organs indicator and the lowest on the mechanism of systemic and pulmonary circulation indicator. These results answer the four research objectives and support the hypothesis that was proposed.

These findings show that turning the flow of blood circulation into a game track helps students understand an abstract process through a concrete, collaborative, and challenging experience. However, because the study used a single-group design without a comparison class, the improvement found cannot yet be fully attributed to the use of Blood Race. The results should be interpreted as preliminary evidence of the feasibility of the product.

The contribution of this study lies in a game design that integrates the double circulation pathway into the board track and in a quantitative evaluation reported comprehensively, from validity and practicality to effectiveness at the indicator and individual levels. With such reporting, teachers and other researchers can judge not only whether the product succeeded but also for which parts of the material it works best.

Science teachers are advised to use Blood Race to reinforce concepts after explaining the material, with heterogeneous groups and a reflection session at the end. Developers are advised to add C3 and C4 quiz cards for the mechanism of circulation indicator and to simplify the rules so that the first round is easier to understand. Future researchers are advised to use a quasi-experimental design with a control class, administer a delayed posttest to measure retention, and involve more schools. Other developers are also advised to provide downloadable print files for the board and cards so that teachers can easily replicate the product.

BIBLIOGRAPHY

- Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*, 45(1), 131–142.
- Arikunto, S. (2013). *Dasar-dasar evaluasi pendidikan* [Fundamentals of educational

- evaluation] (2nd ed.). Jakarta, Indonesia: Bumi Aksara.
- Arsyad, A. (2017). *Media pembelajaran* [Learning media]. Jakarta, Indonesia: Rajawali Pers.
- Branch, R. M. (2009). *Instructional design: The ADDIE approach*. New York, NY: Springer.
- Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. *Review of Educational Research*, 86(1), 79–122. <https://doi.org/10.3102/0034654315582065>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Fajarianti, A. C. (2023). Pengembangan media game edukatif berbasis Wordwall mata pelajaran IPA materi sistem peredaran darah manusia kelas V sekolah dasar [Development of Wordwall-based educational game media for science on the human circulatory system for fifth-grade elementary school]. *Jurnal Penelitian Pendidikan Guru Sekolah Dasar*, 11(1), 113–123.
- Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66(1), 64–74.
- Hamari, J., Shernoff, D. J., Rowe, E., Coller, B., Asbell-Clarke, J., & Edwards, T. (2016). Challenging games help students learn: An empirical study on engagement, flow and immersion in game-based learning. *Computers in Human Behavior*, 54, 170–179.
- Hapsari, N. F., & Mintohari. (2025). Pengembangan media pembelajaran truth or dare board games berbasis air journey dalam pembelajaran IPA sistem pernapasan manusia pada siswa kelas V sekolah dasar [Development of truth or dare board game learning media based on air journey in science learning on the human respiratory system for fifth-grade elementary school students]. *Jurnal Penelitian Pendidikan Guru Sekolah Dasar*, 13(1), 77–91.
- Hardiyanti, Mustami, M. K., & Mu'nisa, A. (2020). Pengembangan game puzzle berbasis Construct 2 sebagai media pembelajaran sistem peredaran darah kelas XI di SMA Negeri 1 Selayar [Development of a Construct 2-based puzzle game as learning media on the circulatory system for eleventh grade at SMA Negeri 1 Selayar]. *Biolearning Journal*, 7(1), 6–11.
- Izza, M., Sukamti, & Winahyu, S. E. (2021). Analisis miskonsepsi materi sistem peredaran darah manusia tema 4 pada siswa kelas V SD [Analysis of misconceptions on the human circulatory system, theme 4, among fifth-grade elementary school students]. *Jurnal Pembelajaran, Bimbingan, dan Pengelolaan Pendidikan*, 1(8), 660–664.
- Munawaroh, A. (2024). Pengembangan media game adventure of blood berbasis role play peredaran darah manusia [Development of the role-play-based adventure of blood game media on the human circulatory system]. *Journal of Education Research*, 5(2), 2431–2441.
- Noda, S., Shiotsuki, K., & Nakao, M. (2019). The effectiveness of intervention with board games: A systematic review. *BioPsychoSocial Medicine*, 13, Article 22. <https://doi.org/10.1186/s13030-019-0164-1>
- Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of game-based learning. *Educational Psychologist*, 50(4), 258–283. <https://doi.org/10.1080/00461520.2015.1122533>
- Sailer, M., & Homner, L. (2020). The gamification of learning: A meta-analysis. *Educational Psychology Review*, 32(1), 77–112. <https://doi.org/10.1007/s10648-019-09498-w>
- Sangi, S. S., & Bata, J. (2023). Pengembangan media pembelajaran berbasis augmented reality pada materi sistem peredaran darah kelas V sekolah dasar [Development of

- augmented reality-based learning media on the circulatory system for fifth-grade elementary school]. *Pendas: Jurnal Ilmiah Pendidikan Dasar*, 8(3).
- Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3-4), 591-611.
- Sugiyono. (2019). *Metode penelitian kuantitatif, kualitatif, dan R&D* [Quantitative, qualitative, and R&D research methods]. Bandung, Indonesia: Alfabeta.