RESEARCH ARTICLE

Development and validation of a conceptual test for the human circulatory system

Emerson G. Gaylan[a,1], Suzette T. Elladora[a,b,2], John Kenneth B. Taneo[a,c,3], Collenn H. Callanga[a,4], Manilyn Piloton-Narca[a,5], Jonavie Becbec[a,6], Joje Mar P. Sanchez[a,d,7,*]

[a]College of Teacher Education, Cebu Normal University, Cebu City, Philippines
[b]College of Education, Cebu Technological University Argao, Municipality of Argao, Philippines
[c]Basic Education Department, University of Cebu-METC, Cebu City, Philippines
[d]Institute for Research in Innovative Instructional Delivery, Cebu Normal University, Cebu City, Philippines

[1]main.15000771@cnu.edu.ph; [2]elladorasuzette@gmail.com; [3]johnkennethtaneo@gmail.com; [4]col.callanga@gmail.com; [5]manilynnarca@gmail.com; [6]jonaviebecbec27@gmail.com; [7]sanchezj@cnu.edu.ph*

For correspondence:
sanchezj@cnu.edu.ph

Article history:
Received: 16 June 2024
Revised: 17 September 2024
Accepted: 1 October 2024
Published: 12 November 2024

10.22219/jpbi.v10i3.34327

p-ISSN: 2442-3750
e-ISSN: 2537-6204

How to cite:
Gaylan, E. G., Elladora, S. T., Taneo, J. K. B., Callanga, C. H., Piloton-Narca, M., Becbec, J., & Sanchez, J. M. P. (2024). Development and validation of a conceptual test for the human circulatory system. *JPBI (Jurnal Pendidikan Biologi Indonesia), 10*(3), 909-919. https://doi.org/10.22219/jpbi.v10i3.34327

Abstract: Current biology assessments often lack a strong emphasis on conceptual understanding, which may limit students' ability to apply knowledge in real-world contexts. This study aimed to develop and validate a concept test to assess high school student's understanding of the human circulatory system. The test was carefully designed using the revised Bloom's Taxonomy and aligned with the K-12 curriculum. Initial content validation by biology education experts ensured accuracy and relevance, leading to refinements in the test items. Item quality was enhanced through dichotomous Rasch analysis following preliminary testing with 100 students. After a second administration to a new group of 100 students, further analysis confirmed the test's reliability and validity. A final round of testing with an additional 100 students yielded a Cronbach's alpha of 0.79, confirming internal consistency. The study concluded that the developed concept test is valid and reliable for assessing students' understanding of the circulatory system, providing teachers with a tool to refine instructional methods. Future recommendations involve expanding and updating test items to ensure relevance, integrating technology-enhanced questions to improve the assessment of student comprehension, and evaluating reliability across diverse contexts.

Keywords: conceptual test; human circulatory system; Rasch analysis; test development; validation

Introduction

The human circulatory system is an essential topic in biology classes (Seah, 2020), and it is necessary to understand how blood, nutrients, gasses, and waste products are moved throughout the body. Students must grasp this idea thoroughly since it provides the foundation for more in-depth research on human physiology and medicine. The heart, blood vessels, and components are all part of the circulatory system and vital to preserving homeostasis (Nilsson & Holmgren, 2021). Despite its complexity, the subject provides abundant educational possibilities to investigate physiological processes, anatomical features, and the integration of systems that support life. In addition to being a fundamental component of biology courses, the human circulatory system is a fascinating subject that can stimulate students' curiosity and encourage a more profound interest in the biological sciences in senior high school (Sele, 2019; Purba et al., 2017).

The human circulatory system has been extensively studied in biology instruction. These studies frequently examine teaching techniques, such as inquiry-based learning (Wulandari et al., 2022; Putra et al., 2018), digital simulations (Heldt et al., 2010; Purba et al., 2017), interactive models (Gnidovec et al., 2020), and conventional lectures. Research aims to find strategies that improve student understanding and engagement. Interactive models allow students to see and control various parts of the circulatory system (Buckley, 2000), and digital simulations may show how dynamic processes like heartbeats and blood flow occur (Mundy et al., 2022). Students are encouraged to investigate issues and conduct experiments as part of inquiry-based learning, promoting deeper understanding via active engagement. Researchers have worked to improve instructional strategies to enhance the effectiveness and engagement of the learning process by analyzing the results of these varied teaching approaches.

Several factors have been looked into when instructing students about the human circulatory system. These include how students view biology, how much they think it is essential to comprehend how the body works, and how they participate in class activities like experiments and discussions (López-Manjón & Angón, 2009; Seah, 2020). Furthermore, factors like study habits, motivation, science process abilities like hypothesizing and experimenting, teacher feedback systems, and students' experiences with the material are all critical (Özen, 2017; Steinmayr et al., 2019). Positive views toward biology, for example, can increase motivation and engagement, and teachers who provide insightful comments can help students learn and clear up misunderstandings by providing students with information about their strengths and weaknesses and helping them develop a sense of self-efficacy (Wilbert et al., 2010). Teachers can better understand the elements influencing students' learning and create ways to address them by researching these variables.

Of these factors, conceptual understanding is the most important when assessing a teaching strategy's effectiveness (Abate et al., 2020; Suskie, 2009). It is the degree of the student's knowledge of the concepts and processes underlying the circulatory system. Conceptual understanding goes beyond simply acquiring information; it involves grasping the underlying principles of how and why a system function as it does, which is a key trait of a scientifically literate individual (Roberts, 2013). For instance, an integrated grasp of anatomy, physiology, and biochemistry is necessary to comprehend the heart's function in pumping blood and the importance of oxygen transport in red blood cells. Rigid learning alone is insufficient in these situations. For students to adequately address associated challenges, including detecting cardiovascular disorders or comprehending how exercise affects heart health, they must have a deeper understanding of the topic.

Though conceptual understanding is acknowledged as a crucial factor, more research is needed concerning the availability of standardized or validated instruments to measure it (Lewis et al., 2015). Although many techniques have been used to assess students' comprehension, very few have undergone extensive reliability and validity testing. Current evaluation methods frequently depend on conventional tests or quizzes, which only partially reflect the breadth of students' conceptual understanding. The lack of reliable assessment tools makes it more challenging to determine how different instructional approaches affect students' conceptual understanding of the circulatory system (Feltovich et al., 2012; Roth, 1990; Black & Wiliam, 2018). As a result, researchers and educators need to be equipped to assess and enhance teaching methods efficiently.

By creating and approving a concept test intended mainly to gauge students' understanding of the human circulatory system, this study sought to close this knowledge gap. This concept test is meant to give researchers and educators a valid and trustworthy instrument for evaluating the efficacy of instructional tactics and enhancing classroom education. Important ideas include the types of blood arteries, the mechanisms of blood circulation, and the structure and function of the heart, which are the main topics of the test (Betts et al., 2022). The study developed a comprehensive and helpful assessment tool by ensuring the test items align with educational standards and are based on scientific evidence. This concept test can aid in a more precise assessment of students' understanding and help teachers improve their strategies to enhance learning results.

The development and validation of this concept test mark a noteworthy contribution to biology and science education in general. Prior studies, including establishing the biology core concept instrument by Cary et al. (2019), effectively constructed instruments for assessing student understanding, but they lacked cross-cultural validation. Although their work improved our understanding of biology misunderstandings, there were issues with its scalability and adaptability to other learning contexts. This study fills these gaps by offering a standardized test designed to assess biological conceptual comprehension, especially concerning the human circulatory system, thus augmenting the capacity to examine and improve instructional strategies by offering a standardized instrument for evaluating conceptual understanding. Ultimately, this could help make biology instruction more successful and evidence-based while ensuring students a thorough and long-lasting grasp of fundamental biological ideas. A validated concept test could also encourage more research in science education by allowing researchers to investigate innovative teaching strategies and how they affect student learning (Gao et al., 2020). This study could fill a gap in biology education and lay the groundwork for future enhancements in teaching and learning in science.

Method

Preparation

The research methodology follows a three-step process proposed by Morales (2012) and Aligway et al. (2024), encompassing the preparation, development, and validation phases. In the preparation phase, careful attention was given to aligning the conceptual test content with the K–12 basic education curriculum. This alignment was achieved through a comprehensive review of curriculum standards, instructional materials, and current lesson plans, ensuring that the test addressed the educational needs and core content areas. During the development phase, the creation of test items was guided by the revised Bloom's Taxonomy, which included a range of cognitive skills from lower to higher-order thinking. This approach ensured that the test items would not only be relevant to the curriculum but would also challenge students across a spectrum of cognitive levels.

Development

In the development stage, test items were crafted based on the revised Bloom's Taxonomy to cover a range of cognitive skills and aligned closely with the curriculum. A 30-question multiple-choice test, with four options per question, was developed to assess students' understanding of the circulatory system. Each question was carefully constructed to meet educational benchmarks and curriculum learning goals. A thorough evaluation of the test material was conducted by six specialists in biology education, including high school educators, university lecturers, and medical professionals. They provided a high content validity rating of 4.61, indicating that the test effectively covers the intended concepts. Content validity is essential to ensure that the instrument accurately measures what it is designed to assess; thus, this high rating supports the test's reliability as a measure of students' comprehension. Expert feedback guided revisions to enhance alignment with educational standards and objectives, further strengthening the test validity and reliability.

Following content validation, the revised test was administered to a preliminary sample of 100 senior high school students. Dichotomous Rasch analysis was conducted to assess item quality, focusing on infit and outfit values, item difficulty, and standard error measures. Items that did not meet the quality criteria specified by Bond and Fox (2007) were revised based on the analysis findings to improve their effectiveness and overall quality.

Validation

The updated concept test was given to an additional 100 SHS students to verify its validity and reliability. The findings were subjected to Aiken's V content validity coefficient analysis to ensure the test evaluated the desired constructs correctly. Item analysis also evaluated item difficulty, discrimination, and distractor effectiveness. Samad's (2004) criteria were employed to assess item difficulty, while both Samad's (2004) and Tamil's (2015) criteria were utilized to evaluate item discrimination. These studies helped shed light on how well each test item performed individually and guided any additional adjustments that were required. Ultimately, a third group of 100 SHS students took the concept test to gauge how reliable the final version was. Reliability coefficients were computed using Cronbach's alpha values; values falling within particular ranges denote different degrees of dependability (Cronbach, 1951). These reliability coefficients offered crucial details regarding the test's internal consistency and capacity to yield reliable findings across several administrations.

Ethical Considerations

Ethical considerations were prioritized throughout the study. All participants provided informed consent, ensuring they understood the purpose and scope of the research. Confidentiality was strictly maintained during data collection to protect participants' privacy and uphold the integrity of the findings. The study protected the rights and welfare of the participants by adhering to ethical norms for the formulation, preparation, and validation of concept tests.

Results and Discussion

Item Distribution

The 30 items of the concept test on the human circulatory system were content validated by the experts in Biology education. These items were distributed according to the learning competency, revised Bloom's cognitive levels, and thinking skill orders, as presented in Table 1.

Table 1. Item distribution of the concept test (N=30)

Competency	Thinking Skill	Cognitive Level	No.	Item
Determine the flow of blood and the cardiac cycle (50.00%)	LOTS (46.67%)	Remembering (10.00%)	1	Identify the vessel carrying oxygen-rich blood
			4	Identify the valves involved in heart sounds
			25	Identify vessels carrying blood away and toward the heart
		Understanding (36.67%)	2	Identify the phase when the heart muscle relaxes and fills
			5	Identify the phase when the AV valve closes
			6	Identify the correct path of blood from the right atrium to the lungs
			10	Arrange organs in order of blood flow from the heart
			14	Explain why atrial walls are thinner than ventricles
			15	Explain oxygen delivery in endotherms with a four-chambered heart
			17	Explain why arteries squirt while veins bleed
			19	Identify true statements about the two circuits
			27	Identify the impact of valve replacement on blood flow
			29	Explain why blood clots are more likely in the lungs than brain
			30	Arrange blood vessels based on decreasing blood pressure
	HOTS (3.33%)	Evaluating (3.33%)	20	Evaluate the role of atria during the cardiac cycle
Discuss how organ systems regulate factors influencing blood pressure (50.00%)	LOTS (20.00%)	Understanding (10.00%)	3	Identify the system influencing blood pressure
			21	Identify true statements about blood pressure
			26	Discuss how organ systems regulate factors influencing blood pressure
		Applying (10.00%)	7	Relate blood vessel diameter to blood flow resistance
			12	Identify appropriate medications for high blood pressure
			28	Identify effective lifestyle changes for regulating blood pressure
	HOTS (30.00%)	Analyzing (13.33%)	8	Describe the interplay between renal and cardiovascular in regulating blood pressure
			11	Identify the impact and intervention for kidney dysfunction on blood pressure
			16	Describe the collaboration between nervous and endocrine systems in blood pressure regulation
			18	Sequence the collaboration between nervous and endocrine systems in blood pressure regulation
		Evaluating (10.00%)	9	Identify effective interventions to reduce high blood pressure
			13	Conclude hydration’s effect on blood flow
			22	Evaluate cardiovascular and musculoskeletal response to exercise
		Creating (6.67%)	23	Design an experiment to investigate hydration and blood flow
			24	Design an experiment to illustrate cardiovascular and renal coordination

The percentage distribution of the concept test on the human circulatory system reflects a purposeful and systematic alignment with the curriculum’s learning goals and the assessment's objectives. While standardized tests often prioritize lower levels of Bloom’s Taxonomy, such as remembering and understanding (Chandio et al., 2016), the developed concept test intentionally balances the evaluation of both lower- and higher-order thinking skills. This approach serves two key objectives: to foster critical thinking and problem-solving abilities in students (Jensen et al., 2014) while also measuring retention of foundational knowledge (Kim et al., 2012). By engaging with the complex physiological processes involved in blood circulation and regulation, students are not only required to recall factual information but also to apply, analyze, and evaluate intricate concepts. This balanced cognitive demand is essential for achieving deep mastery of the subject matter.

The concept test on the human circulatory system's item distribution thoroughly evaluates students' comprehension and application of essential ideas about blood flow, the cardiac cycle, and variables affecting blood pressure. Teachers should focus on understanding and assessing students' understanding and thinking processes. At the same time, they are engaged in learning activities that are meaningful and relevant to the lesson (Lestari et al., 2019). Since it is organized across many cognitive levels, thinking skills, and learning capabilities, this distribution comprehensively assesses students' knowledge and skills. First, the concept test has two primary learning competencies: identifying blood flow and the cardiac cycle and discussing how organ systems control factors affecting blood pressure. These skills cover the essential elements of the circulatory system and its control, guaranteeing that students are evaluated on a wide range of information pertinent to this subject.

Moreover, the order of thinking skills is used to better categorize the items under each learning capability. Recalling and comprehending are examples of lower-order thinking skills (LOTS), which concentrate on fundamental knowledge and grasp of ideas. Tasks such as blood vessel identification, cardiac cycle phases, or truth claims regarding blood pressure are used to evaluate these abilities. However, students must engage in deeper cognitive processes when using higher-order thinking skills (HOTS), which include analyzing, evaluating, and creating. Teachers' exams focus on factual information because students will not exert effort towards deeper learning (Crowe et al., 2008). In addition, when the tests require HOTS but the classes only focus on low-level thinking skills, the students will perform poorly due to a lack of exposure to deeper thinking (Retnawati et al., 2018; Avargil et al., 2011). For example, they must analyze how various organ systems interact to regulate blood pressure or design experiments to examine physiological phenomena.

Furthermore, the updated Bloom's cognitive levels further define each item's level of cognitive processing complexity. While items in the categories of analyzing, evaluating, and creating require higher levels of mental engagement, such as critical thinking, synthesis, and application of knowledge in novel contexts, items in the remembering and understanding category concentrate on recollecting facts and understanding key concepts. Assessing students' learning using Bloom's taxonomy can determine if the goals and desired outcomes are attained. Students who are taught based on the taxonomy often lead to positive outcomes and achievement of lesson objectives (Sudirtha et al., 2022).

Rasch Analysis

The results of the first administration of the concept tests were subjected to dichotomous Rasch analysis to check on the quality of the test items. Table 2 provides valuable insights into the quality of the test items, eventually guiding decisions on whether to accept, revise, or even remove them from the test.

Table 2. Results of the dichotomous Rasch analysis

Item	Proportion	Measure	SE Measure	Infit	Outfit
1	0.70	-0.88	0.29	0.91	0.88
2	0.45	0.21	0.27	1.13	1.15
3	0.50	0.00	0.26	1.00	1.00
4	0.70	-0.88	0.29	0.98	0.97
5	0.17	1.68	0.35	0.99	1.04
6	0.70	-0.88	0.29	0.91	0.88
7	0.33	0.73	0.28	0.97	0.96
8	0.65	-0.65	0.28	0.99	0.98
9	0.90	-2.27	0.43	1.00	1.00
10	0.47	0.14	0.27	1.00	1.00
11	0.52	-0.07	0.26	1.00	1.00
12	0.02	4.18	1.01	1.01	1.38
13	0.42	0.36	0.27	0.97	0.97
14	0.38	0.50	0.27	1.03	1.03
15	0.32	0.81	0.28	1.04	1.06
16	0.20	1.45	0.33	1.01	1.05
17	0.15	1.81	0.37	1.02	1.02
18	0.53	-0.14	0.27	1.10	1.10
19	0.65	-0.65	0.28	1.00	1.01
20	0.33	0.73	0.28	1.02	1.04
21	0.45	0.21	0.27	1.00	0.99
22	0.07	2.73	0.52	1.01	1.01
23	0.77	-1.24	0.31	0.99	1.00
24	0.37	0.58	0.27	0.95	0.94
25	0.37	0.58	0.27	1.02	1.02
26	0.73	-1.05	0.30	1.04	1.06
27	0.47	0.14	0.27	0.91	0.90
28	0.55	-0.21	0.27	1.07	1.08
29	0.45	0.21	0.27	0.98	0.98
30	0.08	2.49	0.47	0.97	0.97

Based on Table 2, none of the items warrants removal from the concept test. Specifically, 20 items were retained due to their appropriate quality contribution to the overall concept test optimization. These items demonstrated proportion scores between 0.33 and 0.90, aligning with Rasch model expectations for an acceptable range of correct responses. The SE Measures for these items were consistently low, ranging from 0.18 to 0.43, indicating precise estimates of difficulty. Furthermore, the infit and outfit values for these items remained near 1.00, which is within the acceptable range for model fit (infit range: 0.93–1.11, outfit range: 0.86–1.00). These indicators suggest that the items perform well within the context of the test, aligning both with the model's expectations and the intended difficulty levels.

Ten items were recommended for revision, such as items 5, 9, 12, 15, 16, 17, 22, 26, 28, and 30. Although the proportion of accurate replies for some of these items is satisfactory, the measurements and fit statistics show deviations from the predicted values. Item 5, for instance, has a low-performance proportion of 17% correct responses, but its measure is high (1.68), indicating low difficulty. Nevertheless, its outfit and infit values are more than 1, suggesting a poor fit with the Rasch model. These differences suggest that these questions may not adequately measure the intended constructs and must be adjusted to increase their validity and reliability. Teachers usually want to have the confidence that their assessments are valuable and will help their students think and reach mastery (Brookhart, 2011).

Content Validity Coefficients

As the Rasch analysis results suggested, the content experts revised and validated ten items again. This time, they assessed the validity of the test items using a five-point agreement. The results of this assessment were subjected to Aiken's V testing using the formula V=S/[n(c-1)]. The test results are shown in Table 3.

Table 3. Results of the Aiken's V test of content validity

Item	Aiken's V	Interpretation
1	1.00	Valid
2	0.92	Valid
3	0.96	Valid
4	1.00	Valid
5	0.88	Valid
6	1.00	Valid
7	0.92	Valid
8	1.00	Valid
9	1.00	Valid
10	0.96	Valid
11	1.00	Valid
12	0.88	Valid
13	0.96	Valid
14	0.92	Valid
15	0.92	Valid
16	0.88	Valid
17	0.88	Valid
18	0.96	Valid
19	1.00	Valid
20	0.96	Valid
21	0.96	Valid
22	0.88	Valid
23	1.00	Valid
24	0.92	Valid
25	0.92	Valid
26	1.00	Valid
27	0.96	Valid
28	0.96	Valid
29	0.92	Valid
30	0.88	Valid

As shown in Table 3, The experts' agreement regarding the content validity of the items is highly agreed upon, as indicated by Aiken's V values, which vary from 0.88 to 1.00. Expert agreement is more robust when Aiken's V values are closer to 1.00, indicating high validity; conversely, values closer to 0.00 indicate lower validity (Aiken, 1985; Rahmawati et al., 2018). All items achieved an Aiken's V value of 0.88 or higher, indicating high-quality items according to the evaluation standards (García-Ceberino et al., 2020). This implies that the items are appropriate and relevant for assessing the conceptual understanding of the human circulatory system. Thus, all items manifest content validity.

Item Difficulty and Discrimination

Since all the items were valid, the concept test was administered to another sample of 100 SHS students. The results of this test administration were analyzed. Table 4 reflects the difficulty and discrimination of the test items. According to the difficulty index in Table 4, most items are classified as moderately complex. This shows that the items are suitable for evaluating the cognitive levels of a wide range of students regarding the human circulatory system because they are both accessible and challenging. In addition, 16.67% of the items are observed to be difficult, while 33.33% of the items are found to be easy. This distribution of difficulty levels ensures that the test adequately challenges the students and allows for some differentiation based on ability.

Moreover, the discrimination index shows that 53.33% of the items show fair discrimination. These items help differentiate students with varying conceptual understanding of the human circulatory system. Furthermore, 3.33% of the items are classed as having excellent discrimination, and the remaining 43.33% are labeled as having good discrimination. This implies that although most items have some degree of discriminatory power among students, some items could have their discriminatory power increased.

Table 4. Difficulty and discrimination indices of the concept test items

Index	Level	No. of Items	Percentage	Overall index
Difficulty	Easy	10	33.33%	0.55 (Moderate)
	Moderate	15	50.00%	
	Difficult	5	16.67%	
Discrimination	Fair	16	53.33%	0.29 (Fair)
	Good	13	43.33%	
	Very Good	1	3.33%	

Nevertheless, all items are accepted in the concept test. This finding means that none of the items were easy or difficult and had negative or poor discriminating power. Therefore, validity is established in the concept test.

Distracter Analysis

After determining the difficulty and discrimination indices, the distracter analysis was conducted. This type of item analysis was done on easy and relatively discriminatory items to check how these distracters effectively function as intended. The distracters should be able to mislead the students toward the correct answer and be plausible enough to distinguish between students who understand and those who do not. Table 5 highlights the analysis of some distracters in the concept tests.

Table 5. Analysis of distracters on some critical items in the concept test

Item	Upper (n=27)				Lower (n=27)				Interpretation
	A	B	C	D	A	B	C	D	
1	**10	0	*17	0	**5	2	*20	*0	A poses a misconception.
11	0	*17	5	**5	3	*8	**13	3	C poses a misconception.
18	*15	3	0	**9	*12	1	2	**12	D poses a misconception.
23	2	0	**3	*22	0	3	**5	*19	C poses a misconception.
25	5	*15	**5	2	3	*12	**7	5	C poses a misconception.

*Correct answer **probable source of misconception

The concept test on the human circulatory system's distracter analysis provides important new information about students' misconceptions, especially those in the lower-performing group. The study focuses on specific issues and shows the percentage of students in the upper (high-performing) and lower (low-performing) groups who choose each choice.

The analysis of student responses revealed several notable misconceptions about the circulatory and related systems. For **Item 1**, the correct answer is **C**(aorta), yet in the top group, ten students incorrectly chose **A** (pulmonary artery), while 17 identified the correct answer. This pattern suggests that option **A** is a source of confusion, as students, particularly high-performing ones, mistakenly believe the pulmonary artery carries oxygen-rich blood from the heart. In reality, the pulmonary artery transports oxygen-poor blood to the lungs. For **Item 11**, which explores the impact of renal failure on the cardiovascular system, the correct response is **B**. Interestingly, five students selected **D**, while 17 top-group students selected **B**, indicating some clarity. However, 13 students in the bottom group opted for **C**, implying a misconception that low sodium levels are directly responsible for low blood pressure and dehydration.

In **Item 18**, which examines the coordination between neurological and endocrine systems in blood pressure regulation, the correct answer is **D**. Nine students in the top group chose this answer correctly. However, 15 mistakenly chose **A**, suggesting that many students misunderstand the nervous system's role in detecting blood pressure changes. This confusion was also evident in the bottom group, where 12 students selected **A**, reflecting a misunderstanding about the initial steps in blood pressure regulation. For **Item 23**, where students were asked to design an experiment related to blood flow and hydration, the correct response is **D**. Although 22 students in the top group correctly selected **D**, three chose **C**.

Similarly, in the bottom group, 19 chose **D**, but five incorrectly opted for **C**. This pattern suggests a misconception that exercise, rather than hydration, is the optimal approach to measuring blood flow velocity, highlighting gaps in understanding experimental design principles.

Lastly, **Item 25** addresses foundational knowledge of blood vessel's roles in transporting blood to and from the heart. The correct answer, **B**, was chosen by 15 students in the top group, with five incorrectly selecting **C**. In the bottom group, 12 selected B, and seven chose **C**. This selection pattern suggests a common misconception: students mistakenly associate capillaries with blood transport to and from the heart rather than recognizing their role in exchanging substances at the tissue level.

In summary, these response patterns reveal specific areas where students hold misconceptions, particularly regarding blood flow dynamics, the role of the pulmonary artery, and experimental design considerations. This analysis emphasizes the importance of targeted instructional strategies to clarify these concepts. The distracter analysis illuminates common misconceptions among students, especially those who scored lower on the test. Items 1, 11, 18, 23, and 25 show distinct patterns in which some erroneous answers are selected repeatedly, pointing to areas of misconception in the human circulatory system.

Reliability Analysis

Since the item analysis results suggest that all items were accepted for the concept test, the test was administered to a third group of 100 SHS students for reliability analysis. The results of the reliability analysis are presented in Table 6.

Table 6. Result of the reliability analysis of the concept test

Variable	Cronbach's alpha	Interpretation
Reliability	0.79	Good

Since the item analysis results suggest that all items were accepted for the concept test, the test was administered to a third group of 100 SHS students for reliability analysis. The results of the reliability analysis are presented in Table 6. With a Cronbach's alpha of 0.79 in Table 6, the concept test's items appear well-correlated, indicating that they most likely measure aspects of the same biology topic. This degree of dependability is appropriate for most educational and psychological tests, meaning the test's measurement is consistent and reliable. According to the interpretation of 0.79 as "good" reliability, relatively higher than the traditional threshold of 0.70 (Cronbach, 1951; Taber, 2017), the test would produce consistent findings if given again under comparable circumstances. The results indicate that the concept test on the human circulatory system significantly advances biology education by helping teachers identify misconceptions and refine instructional strategies. By integrating this tool into the curriculum, student engagement and comprehension are expected to increase.

Furthermore, the test introduces new opportunities for scientific education research and innovation. Its adoption encourages the exploration of varied pedagogical methods and supports continuous improvement in teaching practices. As more concept tests are developed, a comprehensive framework for assessing student understanding will emerge. Ultimately, this concept test is not only a valuable assessment tool but also marks a shift toward evidence-based biology education. Its potential to provide deep insights into student learning makes it an indispensable tool for educators seeking to enhance scientific instruction.

Conclusion

A comprehensive and reliable concept test was developed and validated to check how well the students understand the human circulatory system. Experts validated the test items' content to ensure they matched the required cognitive ability, thinking skills, and learning abilities. The quality of the test items was validated using the dichotomous Rasch analysis, which found that only a few questions needed changes and none needed to be removed. The results of Aiken's V values showed high content validity scores confirming that the questions were appropriate and relevant. The difficulty and discrimination indices ensured the test's ability to distinguish between different levels of student knowledge, which showed a balanced distribution of item difficulty and a fair discrimination power. The reliability analysis validated the test's dependability and internal consistency.

The validated concept test has critical real-world implications for both teachers and students. This test can be a useful diagnostic tool to discover students' misconceptions and knowledge gaps about the human circulatory system. Teachers can adjust their lessons to target students' limitations by identifying the areas in which they struggle, ultimately improving students' performance and comprehension. Additionally, the test can be utilized for formative evaluation, which gives students continuous feedback and directs their study. It is also a valuable summative assessment tool, gauging biology students' academic progress. Future directions for this study could focus on expanding and refining the test items.

While the test has been carefully developed, further adjustments may be necessary as new research emerges and feedback from classroom implementation is gathered. To build on this, subsequent studies could investigate administering the test across various academic settings to validate its reliability and effectiveness among diverse student groups. Additionally, incorporating advanced technological elements, such as interactive simulations or multimedia components, could enhance the assessment process and provide deeper insights into students' understanding. Exploring the test's long-term impact on student learning outcomes and knowledge retention in biology education would also help ensure it remains a valuable resource for teachers and students in understanding the human circulatory system.

Acknowledgment

The authors thank the experts and SHS students who participated in developing and validating the concept test for the human circulatory system.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Author Contributions

All authors contributed equally to the conceptualization, data gathering, analysis, writing, and paper completion.

References

Abate, N. T., Woldeamanuel, Y. W., & Berhane, D. E. (2020). Effectiveness of concept mapping based teaching methods on Grade eight students' conceptual understanding of photosynthesis at Ewket Fana Primary School, Bahir Dar, Ethiopia. *Eurasia Journal of Mathematics Science and Technology Education, 16*(12), em1918. https://doi.org/10.29333/ejmste/9276

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*, *45*(1), 131–142. https://doi.org/10.1177/0013164485451012

Aligway, G. J. B., Delos Angeles, J. C., Collano, A. V., Barroca, E. P., Aves, A. C. D., Catubay, J. F., Edjec, J. T., Butaya, M. D. A., & Cortes, S. T. (2024). Validity and reliability of concept inventory test in human physiology. *JPBI (Jurnal Pendidikan Biologi Indonesia)*, *10*(1), 273-282. https://doi.org/10.22219/jpbi.v10i1.29558

Avargil, S., Herscovitz, O., & Dori, Y. J. (2011). Teaching thinking skills in context-based learning: teachers' challenges and assessment knowledge. *Journal of Science Education and Technology, 21*(2), 207–225. https://doi.org/10.1007/s10956-011-9302-7

Betts, J. G., Young, K. A., Wise, J. A., Johnson, E., Poe, B., Kruse, D. H., Korol, O., Johnson, J. E., Womble, M., & DeSaix, P. (2022, April 20). Ch. 20 Introduction - Anatomy and Physiology 2E | OpenStax. https://openstax.org/books/anatomy-and-physiology-2e/pages/20-introduction

Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education Principles Policy and Practice, 25*(6), 551–575. https://doi.org/10.1080/0969594x.2018.1441807

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences (2nd Edition).* Lawrence Erlbaum Associates Publishers. https://doi.org/10.4324/9781315814698

Brookhart, S. M. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement Issues and Practice, 30*(1), 3–12. https://doi.org/10.1111/j.1745-3992.2010.00195.x

Buckley, B. C. (2000). Interactive multimedia and model-based learning in biology. *International Journal of Science Education, 22*(9), 895–935. https://doi.org/10.1080/095006900416848

Cary, T. L., Wienhold, C. J., & Branchaw, J. (2019). A Biology Core Concept Instrument (BCCI) to teach and assess student conceptual understanding. *CBE—Life Sciences Education*, *18*(3), ar46. https://doi.org/10.1187/cbe.18-09-0192

Chandio, M. T., Pandhiani, S. M., & Iqbal, R. (2016). Bloom's taxonomy: Improving assessment and teaching-learning process. *Journal of Education and Educational Development, 3*(2), 203–221. https://doi.org/10.22555/joeed.v3i2.1034

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334. https://doi.org/10.1007/bf02310555

Crowe, A., Dirks, C., & Wenderoth, M. P. (2008). Biology in Bloom: Implementing Bloom's Taxonomy to enhance student learning in Biology. *CBE Life Sciences Education*, *7*(4), 368–381. https://doi.org/10.1187/cbe.08-05-0024

Feltovich, P. J., Spiro, R. J., & Coulson, R. L. (2012). Learning teaching,and testing for complex conceptual understanding. In N. Frederiksen, R. J. Mislevy, & I. I. Bejar (Eds.), *Test Theory for A New Generation of Tests* (1st ed., pp. 193–230). Routledge. https://doi.org/10.4324/9780203052358-11

Gao, X., Li, P., Shen, J., & Sun, H. (2020). Reviewing assessment of student learning in interdisciplinary STEM education. *International Journal of STEM Education*, *7*(1). https://doi.org/10.1186/s40594-020-00225-4

García-Ceberino, J. M., Antúnez, A., Ibáñez, S. J., & Feu, S. (2020). Design and validation of the instrument for the measurement of learning and performance in football. *International Journal of Environmental Research and Public Health*, *17*(13), 4629. https://doi.org/10.3390/ijerph17134629

Gnidovec, T., Žemlja, M., Dolenec, A., & Torkar, G. (2020). Using augmented reality and the structure–behavior–function model to teach lower secondary school students about the human circulatory system. *Journal of Science Education and Technology, 29*(6), 774–784. https://doi.org/10.1007/s10956-020-09850-8

Heldt, T., Mukkamala, R., Moody, G. B., & Mark, R. G. (2010). CVSIM: An open-source cardiovascular simulator for teaching and research. *Open Pacing, Electrophysiology, & Therapy Journal, 3*, 45–54. https://pubmed.ncbi.nlm.nih.gov/21949555

Jensen, J. L., McDaniel, M. A., Woodard, S. M., & Kummer, T. A. (2014). Teaching to the test. . .or testing to teach: Exams requiring higher order thinking skills encourage greater conceptual understanding. *Educational Psychology Review, 26*(2), 307–329. https://doi.org/10.1007/s10648-013-9248-9

Kim, M., Patel, R. A., Uchizono, J. A., & Beck, L. (2012). Incorporation of Bloom's Taxonomy into Multiple-Choice Examination Questions for a Pharmacotherapeutics Course. *American Journal of Pharmaceutical Education, 76*(6), 114. https://doi.org/10.5688/ajpe766114

Lestari, N. W., Selvia, N. F., & Layliyyah, N. R. (2019). Open-ended approach to students' metacognitive ability. *At-Ta`Lim Journal of Education*, *5*(2), 93–106. https://doi.org/10.36835/attalim.v5i2.263

Lewis, C. C., Fischer, S., Weiner, B. J., Stanick, C., Kim, M., & Martinez, R. G. (2015). Outcomes for implementation science: An enhanced systematic review of instruments using evidence-based rating criteria. *Implementation Science, 10*(1). https://doi.org/10.1186/s13012-015-0342-x

López-Manjón, A., & Angón, Y. P. (2009). Representations of the human circulatory system. *Journal of Biological Education, 43*(4), 159–163. https://doi.org/10.1080/00219266.2009.9656176

Morales, M. P. E. (2012). Development and validation of a concept test in introductory physics for biology students. *The Manila Journal of Science*, 7(2), 26-44. https://www.dlsu.edu.ph/wp-content/uploads/pdf/research/journals/mjs/MJS07-2-2012/MJS07-2-4-morales.pdf

Mundy, N., Poyade, M., & Brownlow, A. (2022). The digital dolphin: Are 3D mobile based and interactive models a useful aid to volunteers on stranding schemes Learning the basic anatomy and pathology of cetaceans? In P. M. Rea (Ed.), *Medical Visualization and Applications of Technology: Vol. IX* (1st ed., pp. 263–295). Springer Cham. https://doi.org/10.1007/978-3-031-06735-8_9

Nilsson, S., & Holmgren, S. (2021). Comparative physiology and evolution of the autonomic nervous system (1st ed). Routledge. https://doi.org/10.1201/9781315139807

Özen, S. O. (2017). The effect of motivation on student achievement. In E. Karadag (Ed.), *The Factors Effecting Student Achievement: Meta-Analysis of Empirical Studies: Vol. VI* (1st ed., pp. 35–56). Springer Cham. https://doi.org/10.1007/978-3-319-56083-0_3

Purba, K. R., Liliana, N., & Kwarrie, Y. N. P. (2017). Development of interactive learning media for simulating human blood circulatory system. *2017 International Conference on Soft Computing, Intelligent System and Information Technology (ICSIIT),* Denpasar, Indonesia, 2017, pp. 275-278**.** https://doi.org/10.1109/icsiit.2017.68

Putra, B. K. B., Prayitno, B. A., & Maridi, M. (2018). The effectiveness of guided inquiry and INSTAD towards students' critical thinking skills on circulatory system materials. *Jurnal Pendidikan IPA Indonesia, 7*(4), 476–482. https://doi.org/10.15294/jpii.v7i4.14302

Rahmawati, R., Rustaman, N. Y., Hamidah, I., & Rusdiana, D. (2018). The development and validation of conceptual knowledge test to evaluate conceptual knowledge of physics prospective teachers on electricity and magnetism topics. *Jurnal Pendidikan IPA Indonesia, 7*(4), 483–490. https://doi.org/10.15294/jpii.v7i4.13490

Retnawati, H., Djidu, H., Kartianom, K., Apino, E., & Anazifa, R. D. (2018). Teacher's knowledge about higher order thinking skills and its learning strategy. *Problems of Education in the 21st Century, 76*(2), 215–230. https://doi.org/10.33225/pec/18.76.215

Roberts, D. A. (2013). Scientific Literacy/Science Literacy. In S. K. Abell, K. Appleton, & D. Hanuscin

(Eds.), *Handbook of Research on Science Education* (1st ed., pp. 743–794). Routledge. https://doi.org/10.4324/9780203824696-32

Roth, K. J. (1990). Developing meaningful conceptual understanding in science. In B. F. Jones & L. Idol (Eds.), *Dimensions of Thinking and Cognitive Instruction* (1st ed., pp. 151–188). Routledge. https://doi.org/10.4324/9780203771686-8

Samad, A. (2004). *Essentials of language testing for Malaysian teachers*. UPM Press http://psasir.upm.edu.my/id/eprint/78579/

Seah, L. H. (2020). What student language reveals about the demands of learning the human circulatory system. *Research in Science Education, 51*(6), 1529–1547. https://doi.org/10.1007/s11165-020-09915-z

Sele, Y. (2019). Optimizing the potential of Children Learning in Science (CLIS) with brain gym: Review on human circulatory concepts. *Biosfer, 12*(2), 238–248. https://doi.org/10.21009/biosferjpb.v12n2.238-248

Steinmayr, R., Weidinger, A. F., Schwinger, M., & Spinath, B. (2019). The importance of students' motivation for their academic achievement – replicating and extending previous findings. *Frontiers in Psychology, 10*. https://doi.org/10.3389/fpsyg.2019.01730

Sudirtha, I. G., Widiana, I. W., & Adijaya, M. A. (2022). The effectiveness of using revised Bloom's Taxonomy-Oriented learning activities to improve students' metacognitive abilities. *Journal of Education and E-learning Research*, *9*(2), 55–61. https://doi.org/10.20448/jeelr.v9i2.3804

Suskie, L. A. (2009). *Assessing student learning: A common sense guide (2nd ed.).* Jossey-Bass. https://www.utep.edu/student-affairs/_files/docs/assessment/assessing-student-learning-suskie-2.pdf

Taber, K. S. (2017). The use of Cronbach's Alpha when developing and reporting research instruments in science education. *Research in Science Education*, *48*(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Tamil, A. M. (2015). Calculating difficulty, discrimination and reliability index/ standard error of measurement. *PPUKM*. https://ppukmdotorg.wordpress.com/2015/04/02/calculating-omr-indexes/

Wilbert, J., Grosche, M., & Gerdes, H. (2010). Effects of evaluative feedback on rate of learning and task motivation: An analogue experiment. *Learning Disabilities: A Contemporary Journal*, *8*(2), 43–52. https://files.eric.ed.gov/fulltext/EJ914126.pdf

Wulandari, D. S., Prayitno, B. A., & Maridi, M. (2022). Developing the guided inquiry-based module on the circulatory system to improve student's critical thinking skills. *Jurnal Pendidikan Biologi Indonesia, 8*(1), 77–85. https://doi.org/10.22219/jpbi.v8i1.16512