RESEARCH ARTICLE

Development and validation of a concept inventory test in photosynthesis for junior high school students

Jessan D. Moradas[a,d,1,*], **Feldene B. Socubas**[b,d,2], **Vhan Marie R. Bacasmas**[c,d,3], **Sheila Gyll L. Arquilita**[d,4], **Sylvester T. Cortes**[d,5]

[a] Bantayan Science High School, Department of Education Region VII, Cebu, Philippines
[b] Sisters of Mary School-Girlstown Inc., City of Talisay, Philippines
[c] University of the Visayas, Cebu City, Philippines
[d] Cebu Technological University, Cebu City, Philippines
[1]jessan.moradas@deped.gov.ph*; [2]denef429@gmail.com; [3]vhanilla1999@gmail.com; [4]gyllnate20@gmail.com; [5]sylvester.cortes@ctu.edu.ph

For correspondence:
jessan.moradas@deped.gov.ph

Article history:
Received: 28 September 2024
Revised: 1 November 2024
Accepted: 7 November 2024
Published: 11 November 2024

10.22219/jpbi.v10i3.36634

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

How to cite:
Moradas, J. D., Socubas, F. B., Bacasmas, V. M. R., Arquilita, S. G. L., & Cortes, S. T. (2024). Development and validation of a concept inventory test in photosynthesis for junior high school students. *JPBI (Jurnal Pendidikan Biologi Indonesia), 10*(3), 898-908. https://doi.org/10.22219/jpbi.v10i3.36634

Abstract: Photosynthesis is a foundational biological topic which often presents learning difficulties because of its complex processes and common misconceptions. Hence, valid and reliable diagnostic tool is needed to identify misconceptions and eventually guide development of instructional interventions. This study aimed to develop a valid and reliable concept inventory test to assess junior high school students' understanding of photosynthesis. It was developed to identify common misconceptions and provide valuable insights into students' cognitive processes. This 20-item concept inventory test crafted with the aid of a table of specifications was based on Revised Bloom's taxonomy of cognitive domain and administered to 355 junior high school students. In validation, results displayed high content validity with a mean of 4.85 and an average Aiken's validity coefficient of 0.96. It also falls in the category moderately difficult with the test difficulty of 0.49, and test discrimination of 0.45. In item classification, 18 items were retained (16 accepted, 2 needs revisions) and while reliability test was employed through Cronbach's alpha (α=0.70) indicating sufficient measure of good and internal consistency. Distractor analysis was also performed to determine possible source of misconception per item. The concept inventory test is found to be a good classroom test with some items to be improved and offers educators a valuable resource for diagnosing student misconceptions and enhancing their teaching practices.

Keywords: assessment and evaluation; biology education; concept inventory test; test development and validation

Introduction

The practice of using a valid and reliable assessment plays an integral part in the teaching-learning process in the field of biology and to all other fields of education. In general, assessment is one of the most critical dimensions of the education process used not only to measure what students learned but also used by educators as feedback mechanism to enhance their teaching practices (Orongan, 2020). It is important that assessment must have a robust framework that ensures validity, reliability, and comprehensive coverage to promote a fair assessment process (Rezai, 2022). Integrative reviews also indicates that assessment does not only measure learning outcomes but can also shape student identities (Nieminen, 2024). Furthermore, a well-developed assessment test provides diagnostic information to the student and teachers regarding common misconceptions of students and other areas of difficulty (Engelhardt, 2009).

Therefore, a good quality assessment is a useful tool for diagnosing conceptual understanding and misconceptions in biology to enhance the quality of teaching-learning process. And one of the strategies that has been utilized in assessing student's level of proficiency of the various concepts in biology is concept inventory test (Cary et al., 2019). A concept inventory is a multiple-choice research-level instrument designed to test conceptual understanding that is usually based on common student misconceptions wherein each question, or item, has one correct answer and a number of incorrect answers, known as distractors (Sands et al., 2018). This instrument is commonly adapted in physics and biology departments (Sikorski & Lee, 2024). In the biological sciences, there has been an increased awareness on the need of concept inventories and developing this assessment tool to drive improved teaching and enhanced students' understanding of key ideas (Klymkowsky & Garvin-Doxas, 2017).

Moreover, as many students fail to understand many important biological concepts, science educators developed concept inventories to assess common alternative conceptions and faulty reasoning of students (D'Avanzo, 2008). In this study, the concept about photosynthesis was chosen for it is a prevalent biology topic that students and teachers often find conceptually challenging (Simmie et al., 2021). Photosynthesis is one of the well documented problematic themes for many young learners have developed misconceptions about it, like plants take their ready-made “food” from the soil. It is one of the most important topics in the biology, as all life on Earth depends on the ability of green plants to produce oxygen and to transform sunlight into chemical energy, which is then used by animals and humans as an energy source (Ahopelto et al., 2011). While photosynthesis is an obligatory part of the science curriculum, research has shown that students often have a poor understanding of it due to common misconceptions (Lim & Poo, 2021; Svandova, 2014; Ray & Beardsley, 2008). Various assessment tools have been developed to diagnose and evaluate secondary and undergraduate students’ conceptual understanding and misconceptions of photosynthesis (e.g., two-tier instrument on photosynthesis (Haslam & Treagust, 1987; Griffard & Wandersee, 2001), Photosynthesis and Plant Respiration Diagnostic Test (PRDT) (Lian & Peng, 2021), close-ended questionnaire (Marmaroti & Galanopoulou, 2006), and diagnostic question cluster (DQC) (Parker et al., 2017), while no concept inventory test (CIT) designed and aligned in K-12 curriculum in the Philippines has been published.

CIT refers to assessment tools useful in evaluating students' understanding regarding core concepts within a subject area (Sands et al., 2018). If a CIT for photosynthesis is developed, it helps identify common misconceptions, such as the following but not limited to the functions of chlorophyll, site of photosynthesis, and distinct processes within it. When misconceptions are identified, teachers can modify their instruction to focus on concepts that need further clarification and track students’ progress leading to curriculum enhancements over time. Hence, this study aims in developing a valid and reliable concept inventory test on the K-12 curriculum learning competency for photosynthesis that can be used as diagnostic, summative or formative assessment as well as a misconception or placement test. The concept inventory test must be developed, validated, and implemented effectively to provide detailed insights about students' cognitive processes (Aligway et al., 2024). The developed concept inventory test is composed of multiple-choice questions which is very efficient in administering and scoring as well as it provides objective evaluation and standardized method of assessment. In addition, the developed concept inventory test focused on assessing Junior High School students in their conceptual knowledge and misconceptions on the topic of photosynthesis.

Method

Design

This research employed a mixed-method approach, using a sequential exploratory design to develop and validate an assessment tool for evaluating junior high school students' conceptual understanding of photosynthesis. The study began with qualitative data collection, including interviews and focus groups, and review of biology curriculum for junior high school student to identify key concepts that should be included in the test. The findings informed the development of the initial test items which were distributed based on Revised Bloom’s Taxonomy. Next, the test undergone quantitative analysis through pilot testing to ensure its reliability and validity. The integration of qualitative and quantitative methods will result in a comprehensive assessment tool tailored to the educational needs of junior high school students.

Respondents

The respondents of this research study include 355 junior high school students (Table 1) with the following inclusion criteria: (a) enrolled in the schools of Cebu Province for S.Y. 2024-2025, (b) ages 13-16 years of age, and (c) junior high school students. The respondents were chosen by purposive sampling, a technique that allowed the researchers to select specific groups on purpose which is also known as selective sampling. The researchers collected data from different schools in Cebu Island. This diverse educational landscape provides a comprehensive setting for gathering varied data points, ensuring a thorough analysis across different types of schools within the community.

Table 1. Profile of the respondents

Grade Level	Gender	
	Male	***Female***
9	46	87
10	92	130
		N=355

Instruments

The instrument utilized in this research is a researcher-designed questionnaire specifically created to assess the validity and reliability of a concept inventory test in Biology for junior high school students. This meticulously crafted questionnaire consists of 20 multiple-choice questions, each offering four answer options: A, B, C, and D. This design intends to evaluate the instrument's effectiveness in measuring the students' understanding of biological concepts. By doing so, the researchers aimed to ensure that the test is a valid and reliable tool for educational assessment in junior high school biology classes. The test questionnaire was designed following a table of specifications, which was structured around the cognitive process dimension of the Revised Bloom's Taxonomy wherein the six categories are Remember, Understand, Apply, Analyze, Evaluate and Create (Krathwohl, 2002).

Data Gathering Procedure

The test's development and validation followed a comprehensive process divided into four key phases: Preparation, Development, Validation, and Assessment. This included three main stages: (i) preparation, (ii) the creation of a concept inventory test in Photosynthesis to evaluate students' concept understanding, and (iii) validation. Each phase was meticulously designed to ensure the test's effectiveness and accuracy.

At the preparation stage, researchers gathered pieces of literature, reviews, observations, and curriculum guide from a Junior High School science subject. These served as their basis for developing their CIT. The researchers meticulously crafted their concept paper and instrument, and with the guidance of their research adviser, their research paper underwent a series of checkpoints. First, the questionnaire underwent a series of content validity from three content experts, and each gave their comments and suggestions to the 20-item CIT. Next, they need to ensure the safety, anonymity, and confidentiality of their respondents. These checkpoints include acquiring a content validity test certificate and ethical clearance.

Once the researchers obtained the clearance, they proceeded to their validation phase, which was to collect their data. Researchers collected the data from the different schools within Cebu Island, once they had identified those schools, they started their validation phase. The researchers asked for the consent of the students during their biology class, and after they gave their permission, the students were informed that they were to be given 20-30 minutes to answer the questionnaire. After the test, the answered questionnaires were collected. The collected questionnaires were tallied, tabulated, and analyzed according to the appropriate data analysis technique.

Data Analysis

The data analysis for this study followed a thorough, multi-step approach. Initially, content validation was carried out using a comprehensive method that included the Table of Specifications, expert judgments, and the Content Validity Index, scores calculation, the checklist from Morales (2012) was utilized for the content validation and is contributed by Aiken's validity test. Following this, data was collected by administering the biology test to a sample of junior high school students, which was then analyzed for reliability and item analysis with a focus on item difficulty and discrimination levels. Distractor analysis was conducted to evaluate the degree to which students selected their responses for each item. To ensure adequate representation, the top 27% of high-scoring students and the bottom 27% of low-scoring students were considered (Morales, 2012). The frequency of students choosing each response was then calculated and recorded for each item.

Ethical Consideration

The researchers adhered to beneficence, respect, and justice throughout the study by minimizing harm, ensuring voluntary participation, and maintaining participant confidentiality. Respondents were fully informed about the study's purpose, their role, and the nature of their commitment, including the 20-30 minutes needed to complete the questionnaire. Participation was voluntary, with no incentives or compensation, and respondents could withdraw at any time. The study involved junior high students from different schools in Cebu Island, selected via purposive sampling, and the researchers ensured minimal emotional risk during data collection. Also, a guardian had co-signed an informed consent due to the participants' minor status, and any conflicts of interest were carefully avoided.

Results and Discussion

Test Development

The 20-item concept inventory test (CIT) in photosynthesis for junior high school students was based on the crafted table of specification (TOS) as reflected in Table 2. The TOS is referred to as the test blueprint that aids teachers align objectives, instruction, and assessment in order to ensure that students are assessed with the right cognitive tasks with balanced test items (Alade & Igbinosa, 2014), hence, alignment must be observed. The learning competency was extracted from the K to 12 DepEd Curriculum guide and was unpacked into five specific objectives. From there, items were distributed according to the level of thinking in the revised bloom's taxonomy wherein 60% of it were classified as lower order thinking skills (LOTS) and 40% for the higher order thinking skills (HOTS).

Table 2. Test Item Allocation Based on the Revised Bloom's Taxonomy

Learning Objectives	**Revised Bloom's Taxonomy** (Krathwohl 2002)					
	Remember	Understand	Apply	Analyze	Evaluate	Create
Enumerate the raw materials needed for photosynthesis	**2** (1,2)	**2** (6,7)				
Identify the importance of the products of photosynthesis	**2** (3,4)	**2** (8,9)	**1** (13)		**1** (18)	
Describe the processes involved in the two stages of photosynthesis	**1** (5)	**2** (10,11)	**2** (14,15)		**1** (19)	
Examine the evidence of photosynthesis		**1** (12)		**2** (16,17)		**1** (20)
Total: 20	5	7	3	2	2	1
	LOTS (60%)			**HOTS** (40%)		

The type of test used was multiple-choice for it can be administered efficiently and at the same time allows for explicit measurement of wide range of knowledge, skills, and competencies (Gierl et al., 2017). Each item in the test consists of a question, known as the stem, and four suggested answers, known as alternatives. Moreover, the alternatives consist of one correct answer, and three plausible but incorrect alternatives known as distractors (Budiyono, 2019).

Content Validation

In order to ensure that the questions in the CIT have content, construct, and face validity, series of validation procedures were followed. The version 1 of the CIT was submitted to three experts in the field of biology for thorough review. There were no items in version that were advised to be rejected or discarded although it has garnered several corrections, comments, and suggestions as shown in Table 3. After all corrections and suggestions were implemented, a version 2 of the CIT was submitted to the same experts for another wave of careful review.

Table 3. Content Validity of Version 1 (v.1) and Version 2 (v.2)

Content Experts	***Version 1***	***Version 2***
Expert 1	**= 4.80** Item 6 – Improve the stem Item 9 – Emphasize other products; not about glucose Item 19 – For improvement; add introductory statement Item 20 – Include diagram with label	**= 4.95** No comment
Expert 2	**= 4.65** Include content, content standard, and learning competency in the test questionnaire.	**= 4.85** No comment
Expert 3	**= 4.60** Very commendable table of specification Item 3 – restate stem into "What is the role of oxygen produced during photosynthesis in the environment?" Item 5 – restate stem into "In what stage of photosynthesis is oxygen produced as by-product?" Item 7 – restate stem into "What do plants do with the glucose created during photosynthesis?"	**= 4.75** No comment
Overall Mean	**= 4.68**	**= 4.85**

Using the 20-item Evaluation Checklist by Morales (2012), all three experts rated version 1 and version. The observed increase (see Table 4) of the overall mean of the version 2 as compared to version 1 was a good indication that in terms of content, construct, and face validity, the test has better quality and ready for administration.

Table 4. Content Validity Coefficient using Aiken's Equation

Checklist Items	Aiken's V (version 1)	Aiken's V (version 2)
1	1.00	1.00
2	1.00	1.00
3	0.83	0.83
4	1.00	1.00
5	0.83	0.83
6	0.75	0.92
7	1.00	1.00
8	1.00	1.00
9	0.75	1.00
10	0.92	0.92
11	1.00	1.00
12	1.00	1.00
13	0.83	0.83
14	0.92	0.92
15	1.00	1.00
16	1.00	1.00
17	1.00	1.00
18	1.00	1.00
19	0.75	1.00
20	0.83	1.00
Average	**0.92**	**0.96**

Furthermore, based on the checklist, Aiken's content validity coefficient was used to authenticate the content validity of the CIT. Coefficient values closer or equal to 1 means that item has more content validity (Aligway et al., 2024). Looking at Table 4 closely, it can be observed that most of the items in version 2 were rated closer or equal to 1 as compared to version 1. In fact, version 2 has greater average content validity coefficient, which is 0.96, few points greater than version 1, which is 0.92. Thus, it is credible to say that experts considered items in version to be more valid in terms of content and construction.

Item Analysis

Item analysis is a process used to assess the quality and effectiveness of each item or question in the test and the test in general. It is of value because it strengthens the skills of the teacher in test construction through providing data which items can be retained and be used again in later exams. Also, it provides data for items that can be improved or revised and the items that needs to be rejected for they have ambiguous or misleading nature (Yahia, 2021). More specifically, conducting item analysis aimed to examine if the item is functioning as intended, assesses the required concepts, discriminates between those who master the content material and those who were not, determines the level of difficulty, and whether the distracters are functioning or not (Ali Rezigalla, 2022).

Item difficulty

Item difficulty refers to the percentage of test takers who provided correct answers for an item. It is of value for it provides ability of an item to determine between test takers who know and who do not know the tested material (Morales, 2012). Item difficulty index determines the level of difficulty of an item, higher value means easier item, lower value means the item is difficult (Aligway et al., 2024).

Table 5. Item Difficulty of Version 2 (v.2)

Item Difficulty Level	Range	Number of Items	Percentage
Very Difficult	0.00 – 0.2	0	0%
Difficult	0.21 – 0.4	7	35%
Moderately Difficult	0.41 – 0.6	7	35%
Easy	0.61 – 0.8	4	20%
Very Easy	0.81 and above	2	10%
Total		20	100%
Test Difficulty		**0.49**	

Table 5 displays the item difficulty levels for Version 2 of a test, categorizing each question based on its difficulty. The "Item Difficulty Level" column includes categories ranging from "Very Difficult" to "Very Easy," with each level defined by a specific range of difficulty index values (0.00 - 0.2 for "Very Difficult" up to 0.81 and above for "Very Easy"). The "Number of Items" column shows how many questions fall within each difficulty range, while the "Percentage" column represents the proportion of questions in each category. In this case, there are no "Very Difficult" items, and most items are "Difficult" (35%) and "Moderately Difficult" (35%), indicating that the test leans towards challenging questions. The average test difficulty is listed as 0.49, suggesting a balanced overall difficulty, where students have a roughly 50% chance of answering each item correctly on average. This distribution of item difficulties helps ensure that the test can effectively assess different levels of student understanding.

Item discrimination

Item discrimination refers to the ability of the item to recognize the students who understand the test material and those who do not. It is an essential technique in the accomplishment of the item analysis for it contributes to the recognition whether items should be discarded or rejected (Gul et al., 2022). Item discrimination is calculated through ranking all test takers from the lowest to the highest scorers. Lower and upper 27% were determined and served as basis for the analysis. The difference in the discrimination index of the upper and lower group determines the discriminating level of an item (Aligway et al., 2024).

Table 6. Item Discrimination of Version 2 (v.2)

Item Discrimination	Range	Number of Items	Percentage
Questionable	-1.00 – -0.60	0	0%
Not Discriminating	-0.59 – -0.09	0	0%
Moderately Discriminating	0.10 – 0.20	0	0%
Discriminating	0.21 – 0.60	17	85%
Very Discriminating	0.61 – 1.00	3	15%
Total		20	100%
Test Discrimination		**0.45**	

Table 6 presents the item discrimination analysis for Version 2 of a test, which assesses the effectiveness of each question in distinguishing between students who understand the material well and those who do not. The "Item Discrimination" column categorizes items by their discrimination power, ranging from "Questionable" (-1.00 to -0.60) to "Very Discriminating" (0.61 to 1.00). "Discriminating" items (0.21 to 0.60) make up the majority, with 17 items (85%), and only 3 items (15%) fall under "Very Discriminating." There are no items categorized as "Questionable," "Not Discriminating," or "Moderately Discriminating," indicating that all items have some level of positive discrimination. The "Test Discrimination" value of 0.45 reflects the overall average discrimination index for the test, suggesting that, on average, the test items are effective in distinguishing between higher and lower-performing students. This distribution shows that most questions are successful at differentiating students based on their understanding, with a small portion of highly effective discriminators. This analysis helps ensure the test accurately measures student knowledge and identifies learning gaps.

From the results of the item difficulty and item discrimination, a decision rule for each item was determined. Decision table (Table 7) presents the items that were accepted, needs revision, rejected, and discarded.

Table 8 summarizes the classification of items from Version 2 of a test based on their performance and effectiveness. Out of the 20 test items, 16 items (80%) are categorized as "accept," meaning they meet the required standards and perform well without needing modification. Two items (10%) are classified as "revise," indicating that they require adjustments to improve clarity or effectiveness but are still generally useful. Another two items (10%) are marked as "reject," suggesting that these items did not perform well and should be removed or replaced in the test. No items are classified as "Discard," meaning all items have some level of potential utility. This classification helps guide test refinement, ensuring that most items contribute to accurate assessment while highlighting a few that need improvement or removal. These data suggest that in the final version (v.3) of the concept inventory test in photosynthesis for junior high school will obtain a total of 18 items which were comprised of 16 accepted items and 2 items for revision.

Table 7. Item Classification derived from Item Analysis of Version 2 (v.2)

Item Number	Level of Difficulty	Discriminating Level	Decision Rule
1	Difficult	Discriminating	Accept
2	Very Easy	Discriminating	Reject
3	Easy	Discriminating	Needs revision
4	Very Easy	Discriminating	Reject
5	Moderately Difficult	Discriminating	Accept
6	Moderately Difficult	Discriminating	Accept
7	Moderately Difficult	Discriminating	Accept
8	Easy	Discriminating	Needs revision
9	Difficult	Discriminating	Accept
10	Moderately Difficult	Discriminating	Accept
11	Difficult	Discriminating	Accept
12	Difficult	Discriminating	Accept
13	Easy	Very Discriminating	Accept
14	Moderately Difficult	Discriminating	Accept
15	Moderately Difficult	Discriminating	Accept
16	Easy	Very Discriminating	Accept
17	Difficult	Discriminating	Accept
18	Moderately Difficult	Very Discriminating	Accept
19	Difficult	Discriminating	Accept
20	Difficult	Discriminating	Accept

Table 8. Summary of Item Classification

	Number of Items	Percentage
Accept	16	80%
Revise	2	10%
Reject	2	10%
Discard	0	0%
Total	20	100%

Reliability

Reliability determines the tests' internal consistency (Morales, 2012). Reliability in testing means that a test consistently measures what it is supposed to, regardless of different conditions. A reliable test produces similar scores no matter when it is taken, which questions are asked, or who scores it, minimizing the impact of random factors or chance (Livingston, 2018). The Statistical Package for Social Sciences (SPSS) software was used as aid to conveniently solve for the value of Cronbach's Alpha.

Table 9 shows that the computed reliability statistics (Cronbach's Alpha) is 0.7 and within the 0.70 – 0.80 which means that the CIT is good for a classroom test as interpreted in Table 10. The alpha reaching 0.7 value signifies sufficient measure of reliability or internal consistency of an instrument which should be a common consideration in science education (Taber, 2018). Furthermore, the result exemplified that although there were few items that needs to be improved, the CIT in photosynthesis for junior high school students has good reliability. It is essential to note that in order to attain higher reliability of coefficient, some items need to be improved or revised and number of items with a high difficulty index should be increased (Ali Rezigalla, 2022).

Table 9. Reliability Statistics for Accepted and Needs Revision Items of Version 2 (v.2)

Number of Items	Cronbach's Alpha
18	0.7

Table 10. Interpretation of Reliability (Source: “SCOREPAK®: Item Analysis”, 2005)

α	Interpretation
≥ 0.50	Questionable reliability. This test should not contribute heavily to the course grade, and it needs revision.
0.50 - 0.60	Suggests need for revision of test, unless it is quite (10 or fewer items). The test definitely needs to be supplemented by other measures (e.g. more tests) for grading.
0.60 - 0.70	Somewhat low. This test needs to be supplemented by other measures (e.g., more tests) to determine grades. There are probably some items which could be improved.
0.70 - 0.80	Good for a classroom test; in the range of most. There are probably a few items which could be improved.
0.80 - 0.90	Very good for a classroom test
≤ 0.90	Excellent reliability, at the level of the best standardized tests

Distractor Analysis

Distractor analysis provides a measure of how well each of the incorrect options contributes to the quality of a multiple-choice item. It helps teachers identify students’ misconceptions, to later on extend guidance in overcoming their errors in thinking and reasoning. Additionally, it allows teachers to determine content areas that needs instructional enhancement and intensive remediation (Gierl et al., 2017).

Table 11 presents the distractor analysis for accepted and needs revision items of the CIT in photosynthesis for junior high school students. In this section, the following data are discussed: (1) items for revision, (2) moderately difficult items, and (3) difficult items.

Table 11. Distractor Analysis for Accepted and Needs Revision Items

Item	Upper (n = 96)				Lower (n = 96)				Remarks
no.	A	B	C	D	A	B	C	D	
1	3	**23	14	*56	8	**53	25	*10	B poses a misconception
3	**14	7	*72	3	**46	10	*25	15	A poses a misconception
5	3	6	*69	**18	25	19	*31	**21	D poses a misconception
6	6	*66	**22	2	21	*23	**22	30	C poses a misconception
7	*66	13	**6	11	*23	27	**35	11	C poses a misconception
8	4	8	*80	**4	13	14	*48	**21	D poses a misconception
9	*63	13	12	**8	*11	20	20	**45	D poses a misconception
10	5	10	**18	*63	29	10	**26	*31	C poses a misconception
11	**31	*49	9	7	**30	*17	14	35	A poses a misconception
12	**30	*32	11	23	**24	*13	33	26	A poses a misconception
13	**5	3	*87	1	**29	20	*27	20	A poses a misconception
14	**13	*63	14	6	**27	*21	28	20	A poses a misconception
15	7	9	**15	*65	25	12	**30	*29	C poses a misconception
16	1	*89	0	**6	23	*28	19	**26	D poses a misconception
17	**23	*60	7	6	**48	*12	24	12	A poses a misconception
18	1	*89	**5	1	9	*22	20	**45	C and D pose misconceptions
19	*35	9	5	**47	*8	31	25	**32	D poses a misconception
20	22	*51	10	**13	26	*14	14	**42	D poses a misconception

*correct answer
**possible source of misconception

In item #3 (see Figure 1), more than 50% of the students from the upper and lower group answered option C, which is the correct answer, reason why the item was tagged easy and despite being discriminating, the item needs revision.

3. What is the role of oxygen produced during photosynthesis in the environment?
    A. It is used by plants for photosynthesis
    B. It helps maintain soil fertility
    C. It is essential for the respiration of most living organisms
    D. It is involved in the decomposition or organic matter

Figure 1. Sample item (#3)

This case is also true for item #8 (see Figure 2) wherein almost 67% of the total number of the upper and lower group chose option C, which is the correct answer, perhaps, the distractors need to be changed by more plausible ones (Ali Rezigalla, 2022).

8. How do plants use the glucose produced in photosynthesis?
    A. For cells division and growth only
    B. To create proteins and nucleic acid only
    C. As an energy source and to build complex carbohydrates like starch and cellulose
    D. Solely for immediate energy needs

Figure 2. Sample item (#8)

In item #6 (Figure 3), around 46% of the students from the upper and lower group answered option B, which is the correct answer, 14% opt for option A, 23% for option C, and 17% for option D. It is easy to point out that option C is the probable source of misconception. It may appear that there was significant difference in the percentage of students who answered correctly and who opt for the incorrect alternatives, but it was also observable that there is proximity of the percentage of students among the incorrect alternatives. This means that all the distractors were functional (Ali Rezigalla, 2022). Hence, the item was tagged as moderately difficult and at the same time discriminating, a good and acceptable index. This item can be saved in the test bank for future use.

6. How do plants intake raw materials of photosynthesis?
    A. Absorbing carbon dioxide into the stem
    B. Water comes through its roots and carbon dioxide then moves to the stomata by diffusion
    C. Water moves to the stomata by diffusion and through its roots and carbon dioxide
    D. Leaves will absorb the water

Figure 3. Sample item (#6)

Another example of item which is moderately difficult, and discriminating is item #7 (Figure 4). There are 46% of the students from the upper and lower group who chose option A, which is the correct answer, while the remaining 54% of the group's responses was distributed with close proximity among the incorrect alternatives: option B with 21%, option C with 22%, and option D with 11%. Hence, the item has good and acceptable indices.

7. What do plants do with the glucose created during photosynthesis?
    A. Convert it into other sugars or store it as starch
    B. Use it to produce more chlorophyll
    C. Convert it into water and carbon dioxide
    D. Release it into the atmosphere

Figure 4. Sample item (#7)

Fascinatingly, unlike the majority of items where the upper and lower group has higher percentage of choosing the correct answer, it was a different case for item #19 (Figure 5). In here, only 22% of the group choose option A, which is the correct answer, strangely, 41% of the group opt for option D, which makes it the probable source of misconception. The remaining 37% was distributed in close proximity to option B and C with 21% and 16% respectively. This item was classified as difficult and discriminating, hence, despite the remarkable distribution of the group's responses, it was still tagged as a good and acceptable item.

19. Parks and gardens can help clean the air. However, cities need more green spaces to fight the rising carbon dioxide levels. A good alternative solution to this problem is enhancing the photosynthetic stages of urban plants. If you are a scientist who can change how plants do photosynthesis in cities, which photosynthetic stage would you enhance and why?
    A. Enhance the Calvin cycle to fix more carbon dioxide
    B. Enhance water absorption to improve overall plant health
    C. Enhance chlorophyll production to better capture light energy
    D. Enhance the light-dependent reactions to produce more oxygen

Figure 5. Sample item (#19)

Item #19 was an item that needed critical thinking skills. The misconception probably lies in the thought that an increase in oxygen in the atmosphere would help fight the rising carbon dioxide, neglecting the fact that to address the rising carbon dioxide, there is a need to hinder its increase in number.
Distractor efficiency is used to evaluate the credibility and functionality of distractors. It can contribute to the acceptance and rejection of items. The items with good distractor efficiency are typically acceptable, conversely, items with poor distractor efficiency, are rejected or needs to be revised (Rezigalla et al., 2019).

Conclusion

The development and validation of the concept inventory test (CIT) in photosynthesis for junior high school students provided a valid, reliable and insightful tool for assessing student understanding and misconceptions in this fundamental biological process. The process of item analysis and validation, including expert reviews and statistical analysis, ensured that the final test version was both reliable and valid. The CIT effectively discriminates between students who have a strong grasp of photosynthesis and those who may hold misconceptions, making it a valuable resource for educators. Future iterations of the CIT should focus on revising items with poor distractor efficiency and enhancing the overall reliability of the test to ensure it serves as a robust diagnostic tool in the educational process.

Acknowledgment

Respectful appreciation to Cebu Technological University for their outstanding assistance and cooperation with research projects. Their dedication to creating a supportive environment for study has been crucial to the accomplishment of our research undertaking. The researchers also extend outmost appreciation to DOST-SEI for providing scholarship STRAND program to reach out and support graduate students in science education.

Conflicts of Interest

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

Author Contributions

J. D. Moradas: review and editing, supervised the crafting of concept inventory test and contributed to written manuscript; **F. B. Socubas:** administered concept inventory test and collected initial data, and contributed to research components and concept inventory test**; V. M. R. Bacasmas:** methodology and contributed to research components and concept inventory test; **S. G. L. Arquilita:** organized the collected data and contributed to research components and concept inventory test; and **S. T. Cortes:** performed data analysis and finalized the manuscript.

References

Ahopelto, I., Mikkilä-Erdmann, M., Anto, E., & Penttinen, M. (2011). Future elementary school teachers' conceptual change concerning photosynthesis. *Scandinavian Journal of Educational Research, 55*(5), 503-515. https://doi.org/10.1080/00313831.2010.550060

Alade, O., & Igbinosa, V. O. (2014). Table of specification and its relevance in educational development assessment. *European Journal of Educational and Development Psychology, 2*(1), 1-17. Retrieved from www.ea-journals.org

Ali Rezigalla, A. (2022). Item analysis: concept and application. *Medical Education for the 21st Century*, 1-17. https://doi.org/10.5772/intechopen.100138

Ali Rezigalla, A., Ibrahim, E., & ElHussein, A. (2019). Item analysis: The impact of distractor efficiency on the discrimination power of multiple choice items. *BMC Medical Education*, 1-7. https://doi.org/10.21203/rs.2.15899/v1

Aligway, G. B., Delos Angeles, J. C., Collano, A. V., Barroca, E. P., Aves, A. D., Catubay, J. F., Edjec, J.T., Butaya, M.D., & Cortes, S. T. (2024). Validity and reliability of a concept inventory test in human physiology. *Jurnal Pendidikan Biologi Indonesia, 10*(1), 265-274. https://doi.org/10.22219/jpbi.v10i1.29558

Budiyono, B. (2019). Five-Option vs Four-Option Multiple-Choice Questions. *IJET (Indonesian Journal of English Teaching), 8*, 1-7. https://doi.org/10.15642/ijet2.2019.8.2.1-7

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

D'Avanzo, C. (2008). Biology Concept Inventories: overview, status, and next steps. *BioScience, 58*(11), 1079–1085. https://doi.org/10.1641/B581111

Engelhardt, P. V. (2009). An Introduction to Classical Test Theory as Applied to Conceptual Multiple-choice Tests: Getting started in physics education research. *American Journal of Physics, 2*(1), 1-

40. https://www.per-central.org/items/detail.cfm?ID=8807
Gierl, M., Bulut, O., Guo, Q., & Zhang, X. (2017). Developing, analyzing, and using distractors for multiple-choice tests in education: A Comprehensive Review. *Review of Educational Research, 87*. https://doi.org/10.3102/0034654317726529
Griffard, P., & Wandersee, J. H. (2001). The two-tier instrument on photosynthesis: What does it diagnose? *International Journal of Science Education, 23*(10), 1039–1052. https://doi.org/10.1080/09500690110038549
Gul, N., Shagufta, S., & Perveen, S. (2022). Item Discrimination of Intelligence Test Items: A quality assurance test for an assessment tool. *3*, 247-253. https://doi.org/10.47205/jdss.2022(3-II)24
Haslam, F., & Treagust, D. F. (1987). Diagnosing secondary students' misconceptions of photosynthesis and respiration in plants using a two-tier multiple choice instrument. *Journal of Biological Education, 21*(3), 203-211. https://doi.org/10.1080/00219266.1987.9654897
Klymkowsky, M. W., & Garvin-Doxas, K. (2017). Understanding randomness and its impact on student learning: Lessons learned from building the Biology Concept Inventory (BCI). *CBE—Life Sciences Education, 7*(2), 227–233. https://doi.org/10.1187/cbe.07-08-0063
Krathwohl, D. R. (2002). A revision of bloom's taxonomy: An overview. *Theory Into Practice, 41*(4), 212-218. https://doi.org/10.1207/s15430421tip4104_2
Lian, L., & Peng, P. Y. (2021). A diagnosis of students' misconceptions of photosynthesis and plant respiration. *Asia Pacific Journal of Educators and Education, 36*(1), 155–176. https://doi.org/10.21315/apjee2021.36.1.9
Lim, H. L., & Poo, Y. P. (2021). A diagnosis of students' misconceptions of photosysthesis and plant respiration. *Asia Pacific Journal of Educators and Education*, *36*(1), 155–176. https://doi.org/10.21315/apjee2021.36.1.9
Livingston, S. C. (2018). Test reliability-basic concepts. Research Memorandum No. RM-18-01. Princeton, NJ: Educational Testing Service.
Marmaroti , P., & Galanopoulou, D. (2006). Pupils' understanding of photosynthesis: A questionnaire for the simultaneous assessment of all aspects. *International Journal of Science Education, 28*(4), 383–403. https://doi.org/10.1080/09500690500277805
Morales, M. P. (2012). Development and validation of a concept test in introductory physics for biology students . *The Manila Journal of Science, 7*(2), 26-44. Retrieved from https://ejournals.ph/article.php?id=2441
Nieminen, J. (2024). How does assessment shape student identities?An integrative review. *Studies in Higher Education*, 1-19. https://doi.org/10.1080/03075079.2024.2334844
Orongan, R. C. (2020). Reliability analysis on teachers' quarterly classroom assessment in basic education. *Liceo Journal of Higher Education Research, 16*(1), 99-107. http://dx.doi.org/10.7828/ljher.v16i1.1370
Parker, J. M., Anderson, C. W., Heidemann, M., Merrill, J., Merritt, B., Richmond, G., & Urban-Lurain, M. (2017). Exploring undergraduates' understanding of photosynthesis using Diagnostic Question Clusters. *CBE-Life Sciences Education, 11*(1), 47-57. https://doi.org/10.1187/cbe.11-07-0054
Ray, A., & Beardsley, P. M. (2008). Overcoming student misconceptions about photosynthesis: A Model and Inquiry-Based Approach using aquatic plants. *Science Activities Classroom Projects and Curriculum Ideas, 45*(1), 13-22. https://doi.org/10.3200/SATS.45.1.13-22
Rezai, A. (2022). Fairness in classroom assessment: development and validation of a questionnaire. *Language Testing in Asia, 12*(1), 17. https://doi.org/10.1186/s40468-022-00162-9
Sands, D., Parker, M., Hedgeland, H., Jordan, S., & Galloway, R. (2018). Using concept inventories to measure understanding. *Higher Education Pedagogies, 3*(1), 60-69. https://doi.org/10.1080/23752696.2018.1433546
SCOREPAK®: Item Analysis. (2005). Office of Educational Assessment, University of Washington. Retrieved March 15, 2012 from http://www.washington.edu/oea/pdfs/resources/item_analysis.pdf
Sikorski , T. R., & Lee, B. K. (2024). Repurposing concept inventories to help students find gaps in explanations. *The Physics Teacher, 62*, 255-258. https://doi.org/10.1119/5.0095315
Simmie, G., Galvin, E., & O'Grady, A. (2021). Alternative concepts in the teaching of photosynthesis: A literature review 2000-2021. *Higher Education of Social Science, 21*(2), 1-11. https://doi.org/10.3968/12304
Svandova, K. (2014). Secondary school students' misconceptions about photosynthesis and plant respiration: Preliminary results. *Eurasia Journal of Mathematics, Science and Technology Education, 10*(1), 59-67. https://doi.org/10.12973/eurasia.2014.1018a
Taber, K. S. (2018). The use of cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education, 48*(1), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2
Yahia, AIO (2021). Post-validation item analysis to assess the validity and reliability of multiple-choice questions at a medical college with an innovative curriculum. *National Medical Journal of India, 34*(6). https://doi.org/10.25259/NMJI_414_20