ORIGINAL ARTICLE

Development of a Two-Tier Diagnostic Test to Assess Misconceptions about Biology Concepts at Primary School

Güntay Tasci

Erzincan Binali Yildirim University, Erzincan, Turkey

Abstract: *Developing tools to identify students' misconceptions about basic biology concepts is necessary. Therefore, a two-tier diagnostic test was developed to determine such misconceptions in primary school (3rd-4th Grade) students. The test content includes two-tiered multiple-choice questions addressing common misconceptions found in the literature on biology subjects in science education at the primary school level. The test's validation and reliability pieces of evidence are described through quantitative and qualitative data analysis of data obtained via the survey method. Data was collected from two samples in two stages, including 74 and 363 primary school students, respectively. The first stage has collected qualitative data by two-tier true-false items based on these common misconceptions. This data was analyzed per qualitative content analysis. Based on these findings, two-tier multiple-choice items were developed, and a two-tier diagnostic test was formed. In the second stage, data obtained was used to quantitative analyze the construct validity, internal consistency, and item parameters of the test. The study's results provided evidence of the validity and reliability of the primary school biology misconceptions diagnostic test (PBMDT), which consists of eight two-tier multiple-choice items.*

Science Insights Education Frontiers 2024; 25(2):4099-4122
DOI: 10.15354/sief.24.or652

How to Cite: Tasci, G. (2024). Development of a two-tier diagnostic test to assess misconceptions about biology concepts at primary school. Science Insights Education Frontiers, 25(2):4099-4122.

Keywords: *Biology Education, Primary School Student, Diagnostic Test, Misconceptions*

About the Author: *Güntay Tasci, Education Faculty, Erzincan Binali Yildirim University, Erzincan, Turkey, E-mail: gtasci@erzincan.edu.tr, ORCID: https://orcid.org/0000-0002-2141-2616*

Correspondence to: *Güntay Tasci at Erzincan Binali Yildirim University in Turkey.*

Conflict of Interests: *None*

Funding: *No funding sources declared.*

AI Declaration: *The author affirms that artificial intelligence did not contribute to the process of preparing the work.*

Introduction

THE CONSTRUCTIVIST approach provides a theoretical framework focusing on prior knowledge, experience, and meaning-making in learning (Fox, 2001). This approach has significantly impacted teaching processes, including field education. As a result, the idea that learners build their concepts related to natural phenomena has become more assertive, particularly in science education (Driver, 1989). Accordingly, Mintzes et al. (2001) reported that meaningful learning occurs when knowledge is structured with appropriately associated concepts to understand natural phenomena. However, in this process, students can also form thoughts unrelated to the scientific meaning of the concepts, resulting in misconceptions (Fisher & Moody, 2002). Although there are different names for misconceptions in the literature, such as preconceptions, naive conceptions, or alternative conceptions, the literature commonly highlights that they occur when learners construct natural phenomena differently from expert opinions or scientific understanding (Coley & Tanner, 2015; Fisher, 1985; Munson, 1994). Various studies tried to identify the sources of misconceptions (Coley & Tanner, 2015, 2012; Hershey, 2004; Yip, 1998). Sadler et al. (2013) suggest that teachers' field knowledge is essential. Yip (1998) points out that teacher characteristic, students' misunderstandings and incomplete understanding of lessons, daily life experiences, and daily language use contribute to forming misconceptions. Hershey (2004) suggests five categories of sources of misconceptions, including "oversimplifications, overgeneralizations, obsolete concepts and terms, misidentifications, and flawed research." Similarly, Tan et al. (2008) emphasize that some rules used in field teaching may become overgeneralizations and lead to misconceptions. Another piece of literature on misconceptions is textbook review studies. Barrass (1984) highlights "misleading terms" in biology textbooks that may create misconceptions. Fisher (1985) highlights the difficulties in changing incorrect textbook information. Based on the literature on the sources of misconceptions, the formation of these misconceptions through course teaching and teaching practices is called "didactic-based misconceptions" (Güngör & Özgür, 2009; Özcan & Bakır, 2023). In addition, Coley & Tanner (2015, 2012) propose a "cognitive construals" framework, analyzing the source of misconceptions with a cognitive psychological approach, and argue that flawed thoughts emerge when learners informally make sense of the world and have identified three categories of thought structures: "teleological, essentialist, and anthropocentric." Researchers expressed this as an effort to create common origins that effectively form different misconceptions in biology. Sadler et al. (2013) reported that studies using qualitative and quantitative techniques to identify misconceptions. In similar studies, researchers have employed various data collection tools such as interviews, drawings, concept maps, concept inventories, and tier diagnostic tests (Gurel et al., 2015; Sesli & Kara, 2012; Tekkaya, 2002; Yip, 1998). Tier tests, one of the diagnostic tests, based on the framework proposed by Treagust (1988), are frequently used in biology education, as in other fields, to investigate misconceptions about different concepts. These tools have been developed with multiple-choice questions in two-tier (Chandrasegaran et al., 2007; Chu et al., 2009; Griffard & Wandersee, 2001; Tsui & Treagust, 2010) in three-tier (Arslan et al., 2012; Caleon & Subramaniam, 2012), and in four-tier (Özden & Yenice, 2017) formats. Every tier of test can have a different number of options. In scoring tier multiple-choice items, it is seen that 0-1 scoring is used for each tier and their combinations (Arslan et al., 2012; Gurel et al., 2015; Ozden & Yenice, 2017; Sesli & Kara, 2012). However, answering multiple-choice items with guessing can be a limitation, which can be reduced by using two-tier tests, although it is essential to consider the potential impact of “diagonal response” when designing these types of tests (Loh et al., 2014). Loh et al. (2014) described “diagonal response” as a situation in which the connections between test tiers affect student responses.

Several studies have identified various biology-related misconceptions in many sub-fields of biology, such as cell biology, ecology, evolution, physiology, biochemistry, and heredity. The participants of these studies were university students, high school students, and secondary school and primary school students (Andrews et al., 2012; Butler et al., 2015; Fisher, 1985; Fisher & Moody, 2002; Kalas et al., 2013; Krall et al., 2009; Meir et al., 2005; Ozay & Haydar, 2003; Öztap et al., 2003; Parker et al., 2012; Sesli & Kara, 2012; Storey, 1992; Tekkaya, 2002; Yip, 1998). In addition, several studies examine the common misconceptions of the textbooks (Gündüz et al., 2016; Storey, 1989, 1990, 1992; Yilmaz et al., 2017). Various studies carried out, especially in science education at the primary school level, reported that misconceptions exist around concepts such as the water cycle, photosynthesis, light and shadows, the concept of living and non-living things, plants, animals, roots, nutrition, and nutrients (Allen, 2014; Asoko, 2002; Dimec & Strgar, 2017; Karpudewan et al., 2017; Lin, 2004; Pine et al., 2001; Topsakal, 2009; Uyanık, 2019). The literature mentioned above emphasizes that educators should identify misconceptions, explain their origins, and eliminate them. The primary school introduces the students to basic terms and factual information, a foundation for scientific concepts. Although primary school science classes do not focus on conceptual knowledge, they are crucial in structuring scientific concepts by providing essential factual information and terms. Misconceptions can significantly hinder scientific understanding, and it is essential to identify and address them early on. Based on this, the study aimed to develop a diagnostic test to determine students’ misconceptions regarding biology concepts included in the primary school curriculum. Considering the various terms in the literature expressing misconceptions, it has become challenging to name misconceptions. However, in literature, it is recommended that educators avoid limiting students' understanding of natural concepts and instead approach their views on concepts flexibly (Maskiewicz & Lineback, 2013; Leonard et al., 2014). Accordingly, the study aimed to identify thoughts that tend to be misconceptions incompatible with scientific understanding in primary school students for different reasons. The current study uses misconceptions for false beliefs, naive concepts, conceptual confusion, alternative concepts, and incorrect conceptualizations that arise in students.

Method

Study Groups

The study involved third- and fourth-grade students from primary school and fifth-grade students from middle school. The students selected per convenience sampling participated voluntarily in the research. The first study sample consists of a total of 74 students. Of these students, 52.7 % (f = 39) are male and 47.3 % (f = 35) are female. The students consist of third-grade (f = 4), fourth-grade (f = 47), and fifth-grade (f = 23) primary school students. The second stage included 362 fourth-grade primary school students, 43.8 % (f = 159) boys and 56.2 % (f = 204) girls.

Development of Data Collection Tools

The stages suggested by Treagust (1988) for developing diagnostic tests were applied in developing the diagnostic test. The primary stages of this process include determining the content, identifying students' misconceptions, and developing the diagnostic test. In this study, the process involved reviewing relevant literature to identify common misconceptions, examining these misconceptions in the first study group, and based on this, developing a two-tier test, which applied to the second study group. The conceptual framework of the test contains the concepts reached by analyzing biology topics in the science curriculum and science textbooks for third and fourth graders. The literature reports common misconceptions about nutrition, foods, living things, the environment, and sensory organs across different education levels (Allen, 2014; Asoko, 2002; Dimec & Strgar, 2017; Karpudewan et al., 2017; Lin, 2004; Pine et al., 2001; Topsakal, 2009; Uyanık, 2019). So, the data collection tool for the first phase included thirty-two-tier true-false items forming from propositions to measure misconceptions first. These true-false items formed in the second tier by the statement "because…," asking students to explain their answers.

Table 1. Distribution of Identified Misconceptions* Across Questions and Tiers.

	First Tier			Second Tier
Misconceptions	Question No	I	II	Options
Sunlight is necessary for germination.	1		+	A
The seed is a non-living thing.	1	+		D E
Soil is a living thing.	1	+		B A
Nutrition and sleep produce our energy. [a]	2		+	A B C E
Foods consist of a single nutritional ingredient.	3	+		A B C E
Foods consist of a single nutritional ingredient.	4			E
Carbohydrates provide immunity.	4	+		C
Proteins do not provide energy.	4	+		B
Only carbohydrates provide energy.	4			A
Vitamins protect our bodies from diseases. [b]	4			D
A balanced diet is getting nutrients rich in protein. [c]	5	+		A B C E
Plants receive food from the soil. [d]	6	+		E
The plant gets energy for life from soil, water, and the sun. [i]	6			B
The flowers are related to growth. [e]	6			C
Sunlight is necessary for germination.	6			D
If a plant has no visible flowers, it is a flowerless plant [f]	7			D
All plants have flowers. [g]	7	+		
The flowers are related to growth.	7			B
If a plant has no flowers visible, it is a flowerless plant.	7			A C
The sensory organ is a source of stimulation.	8			E
Between sensory organs and the brain, there is no connection.	8		+	A B
A sensory organ perceives all things. [h]	8			C
Natural resources are unlimited.	9			A
Recycling produces natural resources.	9	+		B C
Recycling is the matter cycle.	9			E
Humans can create a natural environment.	10	+		A B C D
The life cycle is not associated with reproduction.	11			D
The life cycle is the life process.	11			C
The life cycle occurs in some animals	11		+	B
The life cycle occurs in plants	11			A

+, A, B, C, D, E: It has the expression containing the relevant misconception.
**: Explanations regarding the identified misconceptions were made based on the first study's findings.*
a: They state that our energy is produced by eating and digesting food. This situation points to students' confusion about the concept of nutrition-respiration.
b: Students state that protection from diseases is the duty of vitamins, and growth is the duty of protein.
c: Students consider consuming only or large amounts of protein necessary for a balanced diet.
d: Since they do not know the difference between organic and inorganic matter, they state that the soil nourishes the plant.
e: They state that all plants have flowers and that if they do not have flowers, they cannot grow. This situation points to students' Growth-Reproduction Concept Confusion.
f: Misunderstanding of terms due to lack of knowledge regarding the classification of plants.
g: The students call plants with distinct flowers, such as daisies and roses, "flowers," and these express the confusion between plant and flower concepts.
h: Inability to understand the limitations of sensory organs.
i: Misunderstanding of "life energy" due to lack of knowledge of cellular respiration and photosynthesis.

By analyzing these items (see Data analysis), participant statements regarding literature-based misconceptions were revealed, and created a list of misconceptions (**Table 1**). The two-tier items developed include these misconceptions under appropriate item roots. The first tier of these items contains two options. In the second tier, there are five options, one correct and four distractors (misconceptions), to explain the reason for the option marked in the first tier. This way, a draft diagnostic test of 11 two-tier multiple-choice items was reached. **Table 1** shows the distribution of the misconceptions measured according to the draft items.

Some items address the same misconception for all options, while others contain different misconceptions regarding the same situation. The study examined draft items for “diagonal response,” which revealed that some choices were related to one or both options in the first tier. Some options incorporated student concepts that could be used in both selections in the first tier. However, it identified that some items of the second tier’s options are related to one of the first tier’s options. Therefore, to strengthen the connection between both options and the items in the second tier, some terms in options were removed for specific items (items 1, 4, 9, and 10). However, this was partially achieved in items 2, 3, 5, 6, 7 and 8. Thus, all second-tier items were used with five options to minimize guessing or chance.

Data Collection

The first data collection stage involved administering a 40-minute session of two-tier true-false type 30 items. In the second study, the diagnostic test, consisting of 11 items and Two-tier multiple-choice items, was administered in a 40-minute session. Both scales include an introductory text at the beginning that introduces the test to the participants. In the second stage, in the two-tier multiple-choice diagnostic test, participants were informed to mark the options they were sure were correct in both tiers. The students’ classroom teachers carried out the data collection process.

Data Analysis

Qualitative Analysis Process

Textual data obtained in the study’s first phase was subjected to content analysis. The incorrect statements in the first batch of answers were initially identified and analyzed for misconceptions. These misconceptions were then compared to the literature for supporting evidence. The situations were then coded based on the expressions used by the students, using *in vivo* coding. The coding process was conducted similarly to the example given. The explanations given by three participants who marked the "True" option for the statement "Soil is alive" have been presented.

> *"Because if it were not alive, trees and plants would not be able to produce vitamins and minerals" (Participant 35, Male, 4th Grade).*
> *"Because it grows flowers, trees, and grass in it." (Participant 52, Male, 4th Grade).*
> *"Because then the grass would not grow and the trees would not grow." (Participant 52, Male, 4th Grade).*

These statements indicate that the participants "true" answered that "soil is alive" because they believed that the soil produces food and sustains other living things. From this, it can be concluded that these students perceive the ecological relationships of the soil as an abiotic component with other living things as the "vitality of the soil." This article has identified the participants' misconceptions about living - and non-living things, and a part of these coded" soil is a living thing" as a code. As another example, the explanations of some participants who chose the wrong option for the statement "A seed is a living thing" are as follows.

> *"Because soil is lifeless unless it comes into contact with water and sunlight." (Participant 1, Female, 4th Grade).*
> *"Because it became not plant." (Participant 6, Male, 4th Grade).*
> *"Because they do not do things like nutrition, respiration, movement, and reaction." (Participant 15, Female, 4th Grade).*
> *"Because it is inanimate, but it comes to life when combined with soil, water and sun." (Participant 52, Male, 4th Grade).*
> *"Because it can grow and become alive after you plant it." (Participant 60, Male, 4th Grade).*

These statements supported the participants' belief that the seed is alive when it turns into a plant. Another explanation for this misconception emphasizes the necessity of soil, water and sun for germination. At the same time, the idea that the seed is non-living has emerged based on the idea that the seed cannot perform some life events such as nutrition and respiration. These situations revealed that these students perceived the dormant seed as an "inanimate thing." In this article, this misconception is coded as "A seed is a non-living thing." The process was carried out for all items, and **Table 1** presents the identified misconceptions from the student statements and their corresponding codes.

Table 2. Two-Tier Multiple-Choice Item Scoring Example.

		Scoring		
Item	Misconception	First Tier	Second Tier	Both Tiers
Living being… I. is soil.	Soil is a living thing.	0		
Because: A. It feeds on water and the sun.	Soil is a living thing.		0	0
B. Creatures such as plants and worms live on the soil.	Soil is a living thing.		0	0
C. The seed is dormant*.	+		1	0
D. The seed becomes alive after germination.	The seed is a non-living thing.		0	0
E. When the seed turns into a plant, it becomes alive.	The seed is a non-living thing.		0	0
Other:				
II. is the seed.	+	1		
Because: A. It feeds on water and the sun.	Sunlight is necessary for germination.		0	0
B. Creatures such as plants and worms live on the soil.	The seed is a non-living thing.		0	0
C. The seed is dormant*.	Proper conceptualization		1	1
D. The seed becomes alive after germination.	The seed is a non-living thing.		0	0
E. When the seed turns into a plant, it becomes alive.	The seed is a non-living thing.		0	0
Other:				

**: In the original version, "sleep statement" was used instead of the term "dormant."*

Table 3. Creating Item Scores.

Levels	First Tier	Second Tier	Both Tiers
Scientific conceptualization	1	1	1
Partial information	0	1	0
Partial information	1	0	0
Lack of knowledge	0	0	0

1: True, 0: False

Scoring Rule for Two-Tier Testing

The response combinations to the two-tiered multiple-choice items were considered separate and combined propositions. **Table 2** displays two sets of options - one correct and one incorrect - for the first tier of the items developed for the diagnostic test.

Table 3 presents a summary of the scoring rules. The first and second tiers give a point for selecting the correct option, while selecting the wrong option results in zero points.

Additionally, combining both tiers creates twelve different answer combinations for each question with the “other” option. If both tiers are correct, a cognitive state suited for scientific conceptualization is achieved. For instance, in the example provided in the table, “Living being II is the seed. Because: C. The seed is in a dormant state.” This combination of responses represents the “scientific (proper) conceptualization” level. Apart from this, there are two different combinations. One is to mark the correct option on one tier and the wrong option on the other and mismark both tiers. Correct answer combinations in one tier and incorrect answers in the other were considered a “partial information” situation. Incorrect marking of both tiers was considered the “lack of knowledge” level.

Validity and Reliability

Quantitative and qualitative analyses were used to assess the psychometric properties of the PBMDT. Experts provided feedback on codes created during data coding for the test’s item development phase. This feedback helped ensure that the test accurately covered the misconceptions and content presented in **Table 1**.

Construct validity, reliability, and item parameters of the diagnostic test were assessed using test scores from a large sample group. The study aimed to establish validity and reliability evidence for the first, second, and both tiers of diagnostic test scores using the framework proposed by Adams and Wieman (2011). Construct validity was determined using factor analysis, which involved exploratory factor analysis, minimum residual estimation method, and Promax, one of the oblique rotation techniques. And eigenvalue and parallel analysis graphs were used to determine the dimensions (Adams & Wieman, 2011; Floyd & Widaman, 1995; Ramlo, 2008). The JASP program was utilized to conduct all analyses.

Reliability analysis was conducted after assessing the unidimensionality of the diagnostic test in terms of its measured structure. This process involved calculating several criteria such as McDonald’s ω, Cronbach’s α, Guttman’s λ_6, KR-20 internal consistency, and the Greatest Lower Bound index, all with a 95 % confidence interval. Additionally, the test’s psychometric properties were evaluated by examining the item difficulty index, item discrimination indices, and Point Biserial Correlation values. Criteria recommended by Fisseni (1997; cited in Bühner 2006, p. 140), George and Mallery (2022), and Haladyna and Rodriguez (2013) were taken into account during this process. Lastly, the qualitative findings of the test were quantified using frequency and rate (%) based on the general total.

Findings

Findings Regarding Psychometric Properties

Results Regarding Construct Validity

The Kaiser-Meyer-Olkin test and Bartlett's test results were first examined among the factor analysis results of the diagnostic test. The KMO values obtained with the KMO test for the data set at this stage (KMO = 0.727 for the first tier; KMO = 0.779 for the second tier; KMO = 0.784 for both tiers) are close to the desired value of 0.80. Bartlett's test results were statistically significant (Chi-square (55) = 921.985 for the first tier; Chi-square (55) = 1227.522 for the second tier; Chi-square (55) = 1568.360 for both tiers). Accordingly, the data set was suitable for factor analysis regarding sample size and correlation. According to the rotated analysis results, the first tier showed a single-factor structure, but the second and both tiers showed a two-dimensional structure. When the factor loadings for the rotated analysis were examined, it was seen that the sixth (0.364), second (0.334), and eleventh (0.23) questions in the first tier did not have the desired factor loadings. In the two-factor structure formed in the other tiers, compatible dimensions that can be considered in line with theoretical expectations regarding subject areas, concepts, or misconceptions did not form. In this context, DY6 (0.364), DY2 (0.334), and DY11 (0.231), because their factor loadings were below 0.40, were removed from the analysis, and the analyses were repeated. It was observed that the KMO values obtained by the KMO test for the data set at this stage (KMO = 0.822 for the first tier; KMO = 0.810 for the second tier; KMO = 0.824 for both tiers) provided the desired value of 0.80 recalculated Bartlett's test results were found to be statistically significant (Chi-square (28) = 653.021 for the first tier; Chi-square (28) = 801.279 for the second tier; Chi-square (28) = 892.179 for both tiers). **Table 4** shows the factor loadings and factor characteristics calculated at this stage. Table 4 shows a one-dimensional structure for all tiers with the eight items in the analysis. Factor loadings of the items vary between the lowest 0.495 and 0.801.

It is seen that the eigenvalues for each tier are higher than one, and the highest is (3.031) for both tiers. It can be seen that the variances explained by the tiers are above 0.30, and both tiers explain the highest variance (0.379).

Reliability and Item Analysis Findings

Reliability indices were calculated for eight diagnostic test items based on each tier's data. **Table 5** presents the results.

Table 4. Factor loadings and Rotated Solution Factor Characteristics.

First Tier		Second Tier		Both Tiers	
items	Factor 1	items	Factor 1	items	Factor 1
DY5	0.698	CS3	0.801	T3	0.755
DY8	0.606	CS7	0.664	T7	0.747
DY7	0.572	CS10	0.599	T10	0.613
DY9	0.548	CS5	0.569	T5	0.609
DY3	0.535	CS9	0.533	T9	0.588
DY1	0.529	CS4	0.521	T4	0.528
DY4	0.517	CS8	0.502	T1	0.525
DY10	0.513	CS1	0.495	T8	0.507
Sum Sq. Loadings	2,579		2,817		3,031
Proportion var.	0.322		0.352		0.379

Table 5. Statistics on Diagnostic Test.

Estimate	First Tier			Second Tier			Both Tiers		
	Point estimate	95 % CI		Point estimate	95 % CI		Point estimate	95 % CI	
		LL	UL		LL	UL		LL	UL
McDonald's ω	0.658	0.605	0.711	0.658	0.605	0.711	0.677	0.627	0.728
Cronbach's α	0.658	0.601	0.708	0.666	0.611	0.715	0.686	0.634	0.732
Guttman's λ6	0.637	0.576	0.697	0.647	0.589	0.706	0.668	0.606	0.729
Greatest Lower Bound	0.716	0.679	0.780	0.723	0.686	0.786	0.744	0.703	0.806
Average interitem correlation	0.194	0.153	0.234	0.199	0.157	0.241	0.215	0.168	0.260
means	4.601	4.391	4.811	2.865	2.661	3.069	2.664	2.459	2.868

The first tier has the highest test mean (4.601) but is the lowest for both tiers (2.664). The average correlation value between the items is 0.215 for both tiers. Three reliability estimate methods based on internal consistency yielded values in the 0.60-0.70 range. However, reliability estimates above 0.70 occur in the upper band within confidence intervals. The test showed internal consistency at a rate higher than 0.70 in all three data sets. In addition, the KR20 reliability index calculated for both tiers scores as 0.661. Item difficulty indices for the first tier varied from 0.366 to 0.738, with an average of 0.575, indicating medium difficulty (0.20-0.80; Bühner, 2006). Difficulty indices for the second tier ranged from 0.176 to 0.603, with an average of 0.358. The difficulty level increased to 0.20 for the

Table 6. Item and Options Analysis Findings.

Q	FT	ST A %	B %	C%	D%	E%	Other%	Total%	Empty%	DI	ID	PBC	sd.
1	I	10.8	8.9	0.3	1.4	1.7	0.3	23.3		0.320	0.438	0.53	0.497
	II	6.1	0.6	17.7	24.9	24.4	0.6	74.2	2.5				
3	I	28.9	22.6	8.0	1.7	0.8	0.6	62.5		0.439	0.174	0.58	0.379
	II	6.1	7.7	5.0	13.5	1.1	1.1	34.4	3.0				
4	I	9.6	6.2	5.1	2.3	0.8	0.8	23.9		0.343	0.333	0.54	0.472
	II	6.8	9.6	33.5	16.1	4.8	1.7	72.7	3.4				
5	I	9.2	2.9	2.3	3.4	-	6.3	24.1		0.398	0.567	0.60	0.496
	II	4.0	6.0	1.4	58.2	2,3	0.9	72.8	2.9				
7	I	2.6	6.1	1.4	9.6	4.3	0.3	24.3		0.450	0.198	0.59	0.399
	II	26.4	1.4	21.2	4.3	18.0	1.4	72.8	2.9				
8	I	4.0	1.2	2.6	53.9	1.2	1.2	64.0		0.321	0.521	0.54	0.500
	II	16.1	4.6	6.6	2.0	3.2	0.3	32.9	3.2				
9	I	5.7	12.2	21.2	2.8	17.3	0.6	59.8		0.350	0.229	0.48	0.421
	II	1.4	3.1	9.3	13.9	7.6	1.7	37.1	3.1				
10	I	30.4	3.9	3.9	18.3	1.1	0.8	58.6		0.364	0.204	0.53	0.403
	II	5.4	7.0	2.0	5.1	17.5	1.4	38.3	3.1				
mean										0.579	0.330	0.548	

Q: Question; FT: First Tier; ST: Second Tier; DI: Discrimination Index; ID: Item Difficulty; PBC: Point Biserial Correlation;

second tier. Similar values occurred in both tiers, and the average difficulty changed to 0.333. The difficulty indices of items 3, 7, 9, and 10 in these tiers are high (< 0.229). Accordingly, although these levels are still at the average difficulty level, it can be seen that the difficulty level has increased; that is, they have approached the value of 0.20. Table 6 gives the option analysis results from these results according to the ticking percentages of the options for the items and the item parameters calculated for both tiers.

It is seen that the distractors worked for the correct answer in the first tier for items 1, 4, 5, 7, and 8. The rate of correct answer combinations in these items seems to follow the item difficulty indices. The correct answer combination for the 5th item is 58.2 %, the item difficulty index is 0.567, the correct answer combination for the 8th item is 53.9 %, and the item difficulty index is 0.521, indicating that the items are of average difficulty. The findings showed that items 3, 9, and 10 had lower rates of correct answer combinations (13.5 %, 13.9 %, and 17.5 %, respectively). Also, certain option combinations led to strong distractors: item 3 (I - A, 28.9 %), item 9 (I - C, 21.2 %), and item 10 (I - A, 30.4 %). These items appear difficult and discriminative based on their discrimination (0.439; 0.350; 0.364) and item difficulty indices (0.174; 0.229; 0.204). Upon examining the scoring patterns of the multiple-choice items within the second tier of the test

Table 7. Comprehension Distribution by Item and Subject Areas.

		Scientific Conceptualization		Lack of Knowledge		Partial Information		Total	
Q	Concept	f	%	f	%	f	%	f	%
1	Plant	82	2.82	92	3.17	189	6.51	363	12.50
3	Foods ingredients	58	2.00	232	7.99	73	2.51	363	12.50
4	Functions of foods	121	4.17	86	2.96	156	5.38	363	12.50
5	Balanced diet	206	7.09	92	3.17	65	2.24	363	12.50
7	Plant	72	2.48	95	3.27	196	6.75	363	12.50
8	Sensory organs	190	6.54	133	4.58	40	1.38	363	12.50
9	Natural Resource	79	2.72	220	7.58	64	2.20	363	12.50
10	Natural and Artificial Environment	67	2.31	222	7.64	74	2.55	363	12.50
	Total	875	30,13	1172	40.36	857	29.51	2904	100

Q: Question;

items, it has been determined that the distractor options included with each item are marked at a rate that varies between 1.9 % and 9.5 %. This finding suggests that these specific options have been chosen less frequently than other options. Furthermore, **Table 6** highlights an additional finding regarding the blanking rates, which indicate that these rates fall below 3.4 %. The response rates regarding the “Other” option offered in the items range between 0.3 % and 1.7 %.

Findings Regarding Comprehension Status

Table 7 presents the distribution of participants’ responses by subject area in the PBMDT.

According to the table, participants responded to the fifth item about balanced nutrition (f = 206; 7.09 %), the eighth item about sensory organs (f = 190; 6.54 %), and the fourth item about the functions of foods (f = 121; 4, 17 %) are items with a high level of scientific conceptualization. Scientific conceptualization rates for other substances vary between 2 % and 2.82 %. When examined in terms of all participants, the level of scientific conceptualization is 30.13 %, while the level of partial knowledge and lack of knowledge is approximately 70 %.

Table 8 presents the distribution of misconceptions regarding subject areas.

Based on the table, the participants have higher rates of misconceptions regarding nutritional content than others. Specifically, they have the following misconceptions: “Foods consist of a single nutritional ingredient.” (f = 325, 12 %), “Proteins do not provide energy” (f = 97, 3 %), and “A balanced diet is getting nutrients rich in protein.” (f = 139, 5 %).

Table 8. Distribution of Misconceptions Determined by the PBMDT.

		SC		LoK		PI		Total	
Subject area	Misconceptions	f	%	f	%	f	%	f	%
Foods Ingredients	Foods consist of a single nutritional ingredient.			235	8.4	90	3.2	325	12
	Proper Conceptualization	58	2.1					58	2
Functions of Food Ingredients	Proper Conceptualization	121	4.3					121	4
	Proteins do not provide energy.			60	2.1	37	1.3	97	3
	Vitamins protect our bodies from diseases.			11		57	2.0	68	2
	Only carbohydrates provide energy.					24	0.9	24	1
	Carbohydrates provide immunity.					18	0.6	18	1
Balanced Diet	Proper Conceptualization	206	7.4					206	7
	A balanced diet is getting nutrients rich in protein.			77	2.8	62	2.2	139	5
Plant	If a plant has no flowers visible, it is a flowerless plant.			52	1.9	189	6.8	241	9
	The seed is a non-living thing.			10	0.4	178	6	188	7
	Proper Conceptualization	154	5.5					154	6
	Soil is a living thing.			79	2.8	4	0.1	83	3
	The flowers are related to growth.			21	0.8	5	0.2	26	1
	Sunlight is necessary for germination.					5	0.2	5	0.2
Natural resource	Recycling produces natural resources.			126	4.5	32	1	158	6
	Recycling is the matter cycle.			62	2.2	26	1	88	3
	Proper Conceptualization	79	2.8					79	3
	Natural resources are unlimited.			20	0.7	5	0.2	25	1
Natural and Artificial Environment	Humans can create a natural environment			215	7.7	72	3	287	10
	Proper Conceptualization	67	2.4					67	2
Sensory Organs	Proper Conceptualization	190	6.8					190	7
	Between sensory organs and the brain, there is no connection.			76	2.7	22	1	98	4
	A sensory organ perceives all things.			25	0.9	9	0.3	34	1
	The sensory organ is a source of stimulation.			12	0.4	4	0.1	16	1
	Total	875	31	1081	39	839	30	2795	100

SC: Scientific Conceptualization; LoK: Lack of Knowledge; PI: Partial Information;

Regarding the misconception that foods consist of a single nutritional content, the participants presented as reasons “oranges consist of vitamin C” (28.9 %), and “oranges are a source of vitamins” (22.6 %) (**Table 6**, Item 3). However, regarding the functions of nutritional contents, students give the option that proteins do not provide energy, as well as the reasons that protein is only related to growth and development and carbohydrates are only related to providing energy. Many misconceptions were identified in the two questions numbered one and seven within the subject area of vitality, especially regarding the concept of plants. Plant and flower conceptual confusion was coded 241 times at a rate of 9 % at the levels of partial knowledge and lack of knowledge. In this question, the reasons given by the learners that some plants consist of only leaves and branches were marked 26 % in option A and 21.2 % in option C. However, between 0.2 % and 7 % of misconceptions arise regarding seeds, germination, reproduction - growth, and plant physiology. The misconception that "the seed is non-living" (f = 188, 7 %) was mainly explained by the distractor that "the seed becomes alive when it turns into a plant." The participants explained that "seeds and soil are alive" by nourishing them with water and sun. This misconception means defining vitality for the soil option. In terms of seeds, it shows that students have developed the wrong idea that sunlight is necessary for germination. Regarding natural resources, some students believe that recycling produces these resources. For example, they think that natural resources are regenerated when paper is recycled (21.2 %) and that recycling is the cycle of natural resources (17.3 %). Regarding the natural and artificial environment, some students believe that humans can create a natural environment by building parks and gardens (30.4 %). Finally, some students fail to understand that the sense organs are connected to the brain (4 %).

Discussion

PBMDT Psychometric Properties

The study evaluated the PBMDT's validity and reliability using qualitative and quantitative methods. As per Tregaust's (1988) diagnostic test development stages, the first step is to define misconceptions mentioned in the literature (Allen, 2014; Asoko, 2002; Dimec & Strgar, 2017; Fisher & Moody, 2002; Krall et al., 2009; Karpudewan et al., 2017; Topsakal, 2009; Uyanık, 2019). After testing these misconceptions in the first group's participants, the findings were used to draft items for the PBMDT. The following two steps ensured the face validity of the PBMDT. Firstly, we created the questions based on the guidelines for writing multiple-choice questions by Haladyna and Rodriguez (2013). Secondly, we had a field expert and a teacher examine the questions. Thus, the suitability of the items in terms of the understandability of their contents was examined through interviews. However, within the scope of the options analysis, the low blanking rate was considered a positive finding regarding the clarity and level of suitability. The low rate of use of the "other" option indicates the adequacy of the distractor pool in the second tier for this sample. In the development of diagnostic tests, applications for the readability of the tests have been included (Arslan et al., 2012). In this study, relatedly, primary school teachers consulted their opinions. However, the items were examined regarding the “diagonal response” situation suggested by Loh et al. (2014). In this context, while the risk of logical clues between tiers can be reduced in some items, this situation has been partially eliminated in some items.

The misconceptions presented in Table 1 were initially formed by the literature. These misconceptions were verified in the first sample of the study. In addition, the specification table, one of the recommended ways to ensure content validity (Adams & Wieman, 2011; Crocker & Algina, 1986), was prepared based on this finding. In similar studies in the literature where tier tests were developed, it is seen that the misconceptions claimed to be measured by the tier test were examined with tables of specifications and propositions (Arslan et al., 2012; Liampa et al., 2019; Lin, 2004; Odom & Barrow, 1995). Expert opinions were evaluated regarding misconceptions of the content presented in Table 1, draft items to measure them, and the targeted structure. These applications have provided qualitative evidence for the construct and content validity of PBMDT. Additionally, in the study, a quantitative data set was reached with the scoring rule following the literature (Chandrasegaran et al., 2007; Kilic & Saglam, 2009; Tan et al., 2002; Liampa et al., 2019; Sesli & Kara, 2012). With the data set created in this way, evidence regarding the construct validity of the PBMDT was obtained. Factor analysis is recommended to provide evidence regarding the construct validity of cognitive tests (Adams & Wieman, 2011; Ramlo, 2008). Although it is reported that there are different opinions regarding the size of factor loadings in the literature regarding the construct validity of cognitive tests, items with appropriate factor loadings (> 0.40) were selected (Floyd & Widaman, 1995; Kline, 2005). The PBMDT, which measures misconceptions regarding different biology subjects, showed a one-dimensional structure.

Attention is drawn to the measured structure model to find evidence regarding reliability, another desired criterion in measurement tools. Yurdugul (2005) emphasizes using different reliability indices according to internal consistency, depending on the characteristics of the measured structure. Reliability indices were calculated using four different methods for the PBMDT, which has items with different factor loadings. The KR20 reliability index, calculated as 0.661 according to the scores of both tiers of the PBMDT, was accepted as the reliability index of the scale. These indices are at acceptable levels within the lower and upper bands of the 95 % confidence intervals determined in the literature (George & Mallery, 2022). In addition, it can be said that the reliability indices obtained for the PBMDT are compatible with tier-test development studies. According to Fisseni (1997; cited in Bühner 2006, p. 140), the discrimination of the PBMDT’s items is at appropriate levels (> 0.30) for all tiers. Based on the findings obtained in the study, options analysis was made based on marking frequencies. The literature indicates that the distractors marked less than 5 % are called “non-functioning distractors” (Gierl et al., 2017; Haladyna & Rodriguez, 2013). Tan et al. (2008) stated that the misconceptions that occurred at a 10 % higher rate were common misconceptions. Accordingly, an option with a rate lower than 10 % in the study may not be used in different studies. This study used a five-option structure to prevent “diagonal response.” However, the finding that the misconceptions identified for the distractors are working distractors is a favorable situation for the test’s psychometric properties. Based on these findings, it was concluded that the PBMDT has appropriate psychometric properties and appropriate validity and reliability evidence is provided.

Misconceptions Identified

The PBMDT showed that students, as misconceptions, believe that foods have only one kind of food ingredient and cannot understand the functions of nutrients and balanced nutrition. These misconceptions are similar to those reported by, who suggested that people believe that “we only eat to gain energy, fats are unhealthy, and only fats contain fat” (Allen (2014). The study confirms students’ misconceptions that foods have only one nutritional component and that fats are unhealthy. Allen (2014) presented the misconception that proteins are used only to gain energy. In addition, the current study describes the misconception that protein is only adequate for growth and development. Accordingly, students have a misconception that essential food ingredients only serve one purpose. The idea that the food ingredients identified by the participants may have only one function can be seen as a result of the essentialist thinking in students suggested by Coley and Tanner (2012). The research shows that students often have misconceptions about seed, flower, germination, plant nutrition, and vitality concepts. Plant-related misconceptions were reported in several studies conducted at different education levels (Allen, 2014; Haslam & Treagust, 1987; Hershey, 2004; Krall et al., 2009; Wynn et al., 2017). For example, Allen (2014) reports the misconceptions regarding the concept of vitality, such as that “fire is alive” and “seed is a non-living thing.” In this study, the misconceptions that “seed is inanimate” and “soil is alive” were determined. The misconception that the soil is alive, which emerged among the participants, is based on the misconception that the soil can be fed by water and sun and can create nutrients and other living things. This situation is compatible with attributing nutritional and reproductive characteristics to the soil within the scope of anthropocentric thought proposed by Coley and Tanner (2012). As for the seed, there is a misconception that it is alive when it turns into a plant. Allen (2014) explains these misconceptions, also identified by Topsakal (2009), with the basic common features of living things and proposes the concept of “dormancy” for scientific conceptualization. However, the students’ view that sunlight is necessary for germination and that the plant receives nutrients from the sun or the soil indicates inappropriate thoughts about plant physiology. Dimec & Strgar (2017) reported that several pieces of research on photosynthesis identified misconceptions such as “plants get their nutrients from their environment, especially from the soil; roots are nutritional organs; sunlight is food for plants; and confusion about respiration.” Krall et al. (2009) reported that sunlight is necessary for seed germination, and misconceptions that the plant obtains nutrients through its roots as non-scientific conceptualizations. Hershey (2004) stresses the impact of obsolete concepts and terms, oversimplification, and overgeneralization on the formation of misconceptions. Yip (1998) reported that one of the sources of learners’ misconceptions was the use of these terms in daily life and spoken language. For example, the fact that the term “flowers” is frequently used interchangeably with “plants” in everyday language makes it difficult to understand this concept in a biological context.

Conclusions

A comprehensive literature review is presented that identifies misconceptions and investigates their causes. The developed PBMDT has identified misconceptions that are in line with different opinions found in these literatures regarding the sources of such misconceptions. Similar misconceptions can be found in textbook reviews and among students at various education levels, including university students (Cakiroglu & Boone, 2002; Dimec & Strgar, 2017; Griffard & Wandersee, 2001; Haslam & Treagust, 1987; Krall et al., 2009; Lin, 2004; Yip, 1998). Misleading terms used in textbooks or didactic teachings may encourage learners to fill in the knowledge gaps with thoughts incompatible with scientific understanding. Depending on the age and education level of the learner, these concepts may be simplified and their scope reduced. However, it is essential to distinguish between the daily usage of language and scientific terminology in our lessons and use terms according to scientific understanding. Therefore, it is crucial to avoid oversimplifying and overgeneralizing the knowledge and terminology of the field of biology when educating students. It is also crucial to fill knowledge gaps with scientifically appropriate ideas and to support students in thinking about biology and natural phenomena. However, teachers must pay attention to not supporting ideas that may lead to misconceptions. Teachers’ determination of these tendencies in their students through PBMDT in their classroom evaluations will positively affect concept teaching in advanced education levels.

Ethics Statement

The participants were presented with an informed consent form before participating. The participants were informed that their participation in the test was entirely voluntary and that they could leave the study at any time. During the data collection, the data were anonymized, and no personal and corporate information was collected or used in the research. This study was conducted with the approval of Erzincan Binali Yıldırım University Human Research Ethics Committee (Protocol Number: 02/15).

References

Adams, W. K., & Wieman, C. E. (2011). Development and validation of instruments to measure learning of expert‐like thinking. International Journal of Science Education, 33(9), 1289-1312. DOI: https://doi.org/10.1080/09500693.2010.512369

Allen, M. (2014). Misconceptions in primary science (Second). McGraw Hill Education Open University Press. DOI: https://doi.org/10.1152/advances.2000.24.1.62

Andrews, T. M., Price, R. M., Mead, L. S., McElhinny, T. L., Thanukos, A., Perez, K. E., Herreid, C. F., Terry, D. R., & Lemons, P. P. (2012). Biology undergraduates' misconceptions about genetic drift. CBE Life Sciences Education, 11(3), 248-259. DOI: https://doi.org/10.1187/cbe.11-12-0107

Arslan, H. O., Cigdemoglu, C., & Moseley, C. (2012). A three-tier diagnostic test to assess pre-service teachers' misconceptions about global warming, greenhouse effect, ozone layer depletion, and acid rain. International Journal of Science Education, 34(11), 1667-1686. DOI: https://doi.org/10.1080/09500693.2012.680618

Asoko, H. (2002). Understanding in primary science developing conceptual understanding in primary science. Cambridge Journal of Education, 32(December 2012), 153-164. DOI: https://doi.org/10.1080/0305764022014752

Barrass, R. (1984). Some misconceptions and misunderstandings perpetuated by teachers and textbooks of biology. Journal of Biological Education, 18(3), 201-206. DOI: https://doi.org/10.1080/00219266.1984.9654636

Bühner, M. (2006). Einführung in die Test- und Fragebogenkonstruktion (2nd ed.). Pearson Studium.

Butler, J., Mooney Simmie, G., & O'Grady, A. (2015). An investigation into the prevalence of ecological misconceptions in upper secondary students and implications for pre-service teacher education. European Journal of Teacher Education, 38(3), 300-319. DOI: https://doi.org/10.1080/02619768.2014.943394

Cakiroglu, J., & Boone, J. W. (2002). Preservice elementary teachers' self-efficacy beliefs and their conceptions of photosynthesis and ınheritance. Journal of Elementary Science Education, 14(1), 1-14.

Chandrasegaran, A. L., Treagust, D. F., & Mocerino, M. (2007). The development of a two-tier multiple-choice diagnostic instrument for evaluating secondary school students' ability to describe and explain chemical reactions using multiple levels of representation. Chemistry Education Research and Practice, 8(3), 293-307. DOI: https://doi.org/10.1039/B7RP90006F

Coley, J. D., & Tanner, K. (2015). Relations between intuitive biological thinking and biological misconceptions in biology majors and nonmajors. Cell Biology Education, 14(1), ar8-ar8. DOI: https://doi.org/10.1187/cbe.14-06-0094

Coley, J. D., & Tanner, K. D. (2012). Common origins of diverse misconceptions: Cognitive principles and the development of biology thinking. CBE Life Sciences Education, 11(3), 209-215. DOI: https://doi.org/10.1187/cbe.12-06-0074

Crocker, L., & Algina, J. (1986). Introduction to Classical and Modern Test Theory. Wadsworth Group- Thomson Learning.

Dimec, S. D., & Strgar, J. (2017). Scientific conceptions of photosynthesis among primary school pupils and student teachers of biology. CEPS Journal, 7(1), 49-68. Available at: http://www.pef.uni-lj.si

Driver, R. (1989). Students' conceptions and the learning of science. International Journal of Science Education, 11(5), 481-490. DOI: https://doi.org/10.1080/0950069890110501

Fisher, K. M. (1985). A misconception in biology: Amino acids and translation. Journal of Research in Science Teaching, 22(1), 53-62. DOI: https://doi.org/10.1002/tea.3660220105

Fisher, K., & Moody, D. (2002). Student misconceptions in biology. In K. M. Fisher, J. H. Wandersee, & D. Moody (Eds.), Mapping biology knowledge (pp. 55-75). Kluwer Academic Publishers. Available at: http://www.springerlink.com/index/k26666337645vl41.pdf

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. Psychological Assessment, 7(3), 286-299. DOI: https://doi.org/10.1037/1040-3590.7.3.286

Fox, R. (2001). Constructivism Examined. Oxford Review of Education, 27(1), 23-35. DOI: https://doi.org/10.1080/3054980020030583

George, D., & Mallery, P. (2022). IBM SPSS Statistics 27 Step by Step A simple guide and reference (7th ed.). Routledge.

Gierl, M. J., 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(6), 1082-1116.

Griffard, P. B., & Wandersee, J. H. (2001). The two-tier instrument on photosynthesis: What does it diagnose? International Journal of Science Education, 23(10), 1039-1052. DOI: https://doi.org/10.1080/09500690110038549

Gündüz, E., Yılmaz, M., & Çimen, O. (2016). The Investigation of the 10th Year Biology Text Book of National Education Ministry (MEB) as Regards to Scientific Concept. Journal of Bayburt Education Faculty, 11(2), 414-430. DOI: https://doi.org/10.17152/gefad.305916

Güngör, B., & Özgür, S. (2009). The Causes of the Fifth Grade Students Misconceptions Originated From Didactic About Digestive System Misconceptions. Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education, 3(2), 149-177.

Gurel, D. K., Eryilmaz, A., & McDermott, L. C. (2015). A review and comparison of diagnostic instruments to identify students' misconceptions in science. Eurasia Journal of Mathematics, Science and Technology Education, 11(5), 989-1008. DOI: https://doi.org/10.12973/eurasia.2015.1369a

Haladyna, T. M., & Rodriguez, M. C. (2013). Developing and validating test items. Taylor & Francis. DOI: https://doi.org/10.4324/9780203850381

Haslam, F., & Treagust, D. (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. DOI: https://doi.org/10.1080/00219266.1987.9654897

Hershey, D. R. (2004). Avoid misconceptions when teaching about plants. Available at: https://files.eric.ed.gov/fulltext/ED501356.pdf

Kalas, P., O'Neill, A., Pollock, C., & Birol, G. (2013). Development of a meiosis concept inventory. CBE Life Sciences Education, 12(4), 655-664. DOI: https://doi.org/10.1187/cbe.12-10-0174

Karpudewan, M., Zain, A. N. M., & Chandrasegaran, A. L. (2017). Misconceptions in Science Education: An Overview. In M. Karpudewan, A. L. Chandrasegaran, & A. N. M. Zain (Eds.), Overcoming Students' Misconceptions in Science (pp. 1-9). Springer. DOI: https://doi.org/10.1007/978-981-10-3437-4

Kilic, D., & Saglam, N. (2009). Development of a two-tier diagnostic test concerning genetics concepts: the study of validity and reliability. Procedia - Social and Behavioral Sciences, 1(1), 2685-2686. DOI: https://doi.org/10.1016/j.sbspro.2009.01.474

Kline, R. B. (2005). Principles and Practice of Structural Equation Modeling (Second). The Guilford Press.

Krall, R. M., Lott, K. H., & Wymer, C. L. (2009). Inservice elementary and middle school teachers' conceptions of photosynthesis and respiration. Journal of Science Teacher Education, 20(1), 41-55. DOI: https://doi.org/10.1007/s10972-008-9117-4

Liampa, V., Malandrakis, G. N., Papadopoulou, P., & Pnevmatikos, D. (2019). Development and evaluation of a three-tier diagnostic test to assess undergraduate primary teachers' understanding of ecological footprint. Research in Science Education, 49(3), 711-736. DOI: https://doi.org/10.1007/s11165-017-9643-1

Lin, S. W. (2004). Development and application of a two-tier diagnostic test for high school students' understanding of flowering plant growth and development. International Journal of Science and Mathematics Education, 2(2), 175-199. DOI: https://doi.org/10.1007/s10763-004-6484-y

Loh, A. S. L., Subramaniam, R., & Tan, K. C. D. (2014). Exploring students' understanding of electrochemical cells using an enhanced two-tier diagnostic instrument. Research in Science and Technological Education, 32(3), 229-250. DOI: https://doi.org/10.1080/02635143.2014.916669

Maskiewicz, A. C., & Lineback, J. E. (2013). Misconceptions are "so yesterday!" CBE Life Sciences Education, 12(3), 352-356. DOI: https://doi.org/10.1187/cbe.13-01-0014

Meir, E., Perry, J., Stal, D., Maruca, S., & Klopfer, E. (2005). How Effective Are Simulated Molecular-level Experiments for Teaching Diffusion and Osmosis? Cell Biology Education, 4(3), 235-248. DOI: https://doi.org/10.1187/cbe.04-09-0049

Mintzes, J. J., Wandersee, J. H., & Novak, J. D. (2001). Assessing understanding in biology. Journal of Biological Education, 35(3), 118-124. DOI: https://doi.org/10.1080/00219266.2001.9655759

Munson, B. H. (1994). Ecological misconceptions. Journal of Environmental Education, 25(4), 30-34. DOI: https://doi.org/10.1080/00958964.1994.9941962

Nehm, R. H. (2019). Biology education research: building integrative frameworks for teaching and learning about living systems. Disciplinary and Interdisciplinary Science Education Research, 1(1). DOI: https://doi.org/10.1186/s43031-019-0017-6

Odom, A. L., & Barrow, L. H. (1995). Development and application of a two - tier diagnostic test measuring college biology students' understanding of diffusion and osmosis after a course of instruction. Journal of Research in Science Teaching, 32(1), 45-61. DOI: https://doi.org/10.1002/tea.3660320106

Ozay, E., & Haydar, O. (2003). Secondary students' interpretations of photosynthesis and plant nutrition. Journal Of Biological Education, 37(2), 68-70.

Özcan, F., & Bakır, S. (2023). Thematic content analysis of studies concerning misconceptions in science education in Turkey. Sakarya University Journal of Education, 13(2), 257-285. DOI: https://doi.org/10.19126/suje.1254983

Özden, B., & Yenice, N. (2017). "Kuvvet ve Enerji" ünitesine yönelik üç aşamalı kavramsal anlama testi geliştirme çalışması. Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi, 11(2), 432-463. DOI: https://doi.org/10.17522/balikesirnef.373421

Öztap, H., Özay, E., & Öztap, F. (2003). Teaching cell division to secondary school students: an investigation of difficulties experienced by Turkish teachers. Journal of Biological Education, 38(1), 13-15. DOI: https://doi.org/10.1080/00219266.2003.9655890

Parker, J. M., Anderson, C. W., Heidemann, M., Merrill, J., Merritt, B., Richmond, G., & Urban-Lurain, M. (2012). Exploring undergraduates' understanding of photosynthesis using diagnostic question clusters. CBE Life Sciences Education, 11(1), 47-57. DOI: https://doi.org/10.1187/cbe.11-07-0054

Pine, K. J., Messer, D., & St. John, K. (2001). Children's misconceptions in primary science: A survey of teachers' views. International Journal of Phytoremediation, 19(1), 79-96. DOI: https://doi.org/10.1080/02635140120046240

Ramlo, S. (2008). Validity and reliability of the force and motion conceptual evaluation. American Journal of Physics, 76(9), 882-886. DOI: https://doi.org/10.1119/1.2952440

Sadler, P. M., Sonnert, G., Coyle, H. P., Cook-Smith, N., & Miller, J. L. (2013). The influence of teachers' knowledge on student learning in middle school physical science classrooms. American Educational Research Journal, 50(5), 1020-1049. DOI: https://doi.org/10.3102/0002831213477680

Sesli, E., & Kara, Y. (2012). Development and application of a two-tier multiple-choice diagnostic test for high school students' understanding of cell division and reproduction. Journal of Biological Education, 46(4), 214-225. DOI: https://doi.org/10.1080/00219266.2012.688849

Storey, R. D. (1989). Misconceptions biology : Cell structure. American Biology Teacher, 52(4), 213-218.

Storey, R. D. (1990). Textbook errors cell in misconcepions metabolism biology. American Biology Teacher, 53(6), 339-343.

Storey, R. D. (1992). Textbook Errors & Misconceptions in Biology: Cell Energetics. American Biology Teacher, 54(3), 161-166. DOI: https://doi.org/10.2307/4449321

Tan, K. C. D., Taber, K. S., Liu, X., Coll, R. K., Lorenzo, M., Li, J., Goh, N. K., & Chia, L. S. (2008). Students' conceptions of ionisation energy: A cross-cultural study. International Journal of Science Education, 30(2), 263-283. DOI: https://doi.org/10.1080/09500690701385258

Tan, K. C. D., Goh, N. K., Chia, L. S., & Treagust, D. F. (2002). Development and application of a two - tier multiple choice diagnostic instrument to assess high school students' understanding of inorganic chemistry qualitative analysis. Journal of Research in Science Teaching, 39 (4), 283-302.

Tekkaya, C. (2002). Misconceptions as barrier to understanding biology. HU Journal of Faculty of Education, 23, 259-266.

Topsakal, Ü. U. (2009). The effectiveness of thematic learning method on

Remedying misconceptions about living and non-living things. Sakarya Üniversitesi Eğitim Fakültesi Dergisi, 17(Mayıs), 219-234.

Treagust, D., F., (1988) Development and use of diagnostic tests to evaluate students' misconceptions in science. International Journal of Science Education, 10 (2), 159-169, DOI: DOI: https://doi.org/10.1080/0950069880100204

Tsui, C. Y., & Treagust, D. (2010). Evaluating secondary students' scientific reasoning in genetics using a two-tier diagnostic instrument. International Journal of Science Education, 32(8), 1073-1098. DOI: https://doi.org/10.1080/09500690902951429

Uyanık, G. (2019). Determination of the misconceptions towards science concepts of primary school students. TÜBAV Bilim, 12(4), 45-54.

Wynn, A. N., Pan, I. L., Rueschhoff, E. E., Herman, M. A. B., & Archer, E. K. (2017). Student misconceptions about plants - a first step in building a teaching resource. Journal of Microbiology & Biology Education, 18(1). DOI: https://doi.org/10.1128/jmbe.v18i1.1253

Yilmaz, M., Ertunç G., Hatun Diken, E., & Çimen, O. (2017). The analysis of biology topics in the 8th grade science textbook in terms of scientific content. Erzincan University Jounal Faculty of Education, 19(3), 17-35.

Yip, D. Y. (1998). Teachers' misconceptions of the circulatory system. Journal of Biological Education, 32(3), 207-215. DOI:

https://doi.org/10.1080/00219266.1998.9655622

Yurdugul, H. (2005). The comparison of reliability coefficients in parallel, tau-equivalent, and congeneric measurements. Ankara University, Journal of Faculty of Educational Sciences, 39(1), 15-37. DOI: https://doi.org/10.1501/egifak_0000000127

Received: June 05, 2024
Revised: July 04, 2024
Accepted: July 04, 2024