J-STEAM

Journal of STEAM Education
Journal of Science, Technology, Engineering, Mathematics and Art Education
2022, July (Issue: 2, Volume: 5)

STEAM Eğitimi Araştırmaları Derneği

Atmosphere-Related Environmental Problems Diagnostic Test: A Validation Study in Bosnia and Herzegovina University Student

Turajlić Nizama[1], Nermin Đapo[2], Mahir Gajević[3], Vanes Mešić[4], Rifat Škrijelj[5]

Article History

Received: 22 Feb 2022

Received in revised form:

5 Apr 2022

Accepted: 28 Apr 2022

Published: 1 July 2022

ABSTRACT

This study was conducted with three aims. The first aim of our study was to examine both construct-related validity and content validity of the Atmosphere-related environmental problems diagnostic test in Bosnia and Herzegovina university student sample. The Atmosphere-related environmental problems diagnostic test is a three-tier multiple-choice diagnostic test consisting of 13 questions on global warming, greenhouse effect, ozone layer depletion and acid rain. The second aim of this study was to examine scientific understanding as well as misunderstanding of atmosphere-related environmental problems among B&H university student sample. Finally, the third aim of our study was to compare scientific understanding and misconceptions of the atmosphere-related environmental problems with respect to educational background. A total of 445 students of three faculty participated in the research. Results indicate that Atmosphere-related environmental problems diagnostic test measures a single construct of general scientific knowledge about atmosphere-related environmental problems. In addition, the content validity and reliability were satisfactory. Results obtained in our study show that students' overall understanding of each content area was low but comparable to knowledge of pre-service teachers in the USA. Similar to earlier research, most incorrect answers resulted from lack of knowledge rather than from misconceptions. Students who attended ecology classes scored higher than students who had not attended these classes. However, although having higher scores on Atmosphere-related environmental problems diagnostic test, students who attended ecology classes also exhibited more misconceptions related to atmosphere-related environmental problems compare to who had not attended ecology classes. This finding indicates that in the context of university education in Bosnia and Herzegovina, one has to also check for possible sources of didaktikogenic misconceptions related to environmental education. Identifying and understanding their possible origins is critical for designing better educational materials and programs.

Keywords: Three-Tier Diagnostic Test**,** Misconception, Environmental Education, University Students

[1]PhD student, University of Sarajevo, Faculty of Science, nizamaturajlic@yahoo.com, 0000-0001-6911-5816
[2]Professor, University of Sarajevo, Faculty of Philosophy, nermin.djapo@ff.unsa.ba, 0000-0001-5957-2589
[3]Professor, University of Sarajevo, Faculty of Science, mahir.gajevic@pmf.unsa.ba, 0000-0001-7478-7940
[4]Professor, University of Sarajevo, Faculty of Science, v.mesic@pmf.unsa.ba, 0000-0003-3337-3471
[5]Professor, University of Sarajevo, Faculty of Science, rskrijelj@pmf.unsa.ba, 0000-0003-0194-2334

Turajlić, N., Đapo, N., Gajević, M., Mešić, V. & Škrijelj, R. (2022). Atmosphere-related environmental problems diagnostic test: A validation study in Bosnia and Herzegovina university student. *Journal of STEAM Education, 5*(2), 117-135. https://doi.org/10.55290/steam.1084655

INTRODUCTION

Over the past decades, various studies were conducted that show us the effect of human activities on climate changes on Earth. Although there are certain regional differences in intensity and volume of climate changes, it is evident that the global sea level rise and rise in surface temperature lead to an increased precipitation and the more extreme tropical storms (Salinger, 2005; United Nation's Intergovernmental Panel on Climate Change (IPCC), 2007). Climate changes are globally one of the most serious problems that threaten our environment with severe social, environmental and economic consequences. Climate changes may disrupt food and water availability, energy production and consumption, plant and animal survival, human health and social and political stability (Paterson, 1996). Climate change is the defining issue that require serious attention of scientists, politicians and the public in general. Climate change prevention requires an urgent international cooperation and consensus, as well as serious changes in the modern lifestyle.

Education is crucial in understanding processes that cause climate change. It also contributes to the scientific understanding of effects and causes of climate change. Compulsory education in Bosnia and Herzegovina, elementary and secondary, offers classes on environmental issues from 5th grade (11-year-old pupils) through the Nature and Society course while these issues are taught in Chemistry, Biology and Geography in the secondary schools. Teaching in the lower grades of elementary schools focuses on acid rain (AR) and ozone layer depletion (OLD). Higher grades offer teaching on global warming causes and effects (GW), greenhouse effect (GE), ozone layer depletion (OLD) and acid rain (AR). High schools continue education on environmental issues and climate change in biology, chemistry and geography. Schools for higher education offer relevant education on climate change in the biology, chemistry and geography departments. Students therefore acquire practical and theoretical knowledge that contribute to the scientific understanding of causes, processes and effects of climate change.

Relatively strong focus on climate change issues in the courses offered throughout the compulsory education suggest that the high school graduates obtain solid knowledge on climate change issues. Studies on the other hand show that many students and teachers have misconceptions on these issues.

Misconceptions on environmental issues had been analyzed through various diagnostic processes. Closed-form Likert type questionnaires (e.g. Boys et al., 1993,1995; Groves & Pugh, 1999, 2002; Michail, Stamou & Stamou, 2007; Pekel & Ozay, 2005), interviews (e.g. Pruneau et al., 2001; Rye et al., 1997; Summers et al., 2000), and open-ended questionnaires (e.g. Anderson & Wallin, 2000; Gowda et al., 1997, Papadimitriou, 2004), are the most common methods in identifying environmental misconceptions.

Multiple choice tests are frequently used in the misconception studies. Although the conventional multiple-choice tests are practical in evaluating content knowledge, test answers do not provide profound understanding of misconceptions. Treagust (1988) developed the two-tier test that requires an explanation for the answer given by the student. The two-tier test cannot differentiate mistakes caused by the lack of knowledge from mistakes caused by the existence of alternative conceptions. Thus, researchers introduced the level of certainty of the respondents for their answers to the first two tiers (Hasan et al., 1999) as a third tier. A right answer with a high degree of confidence indicates a thorough understanding of the related concept. However, answers with low confidence are considered as a lack of knowledge irrespective of if the answer is correct or wrong. A wrong answer accompanied by a strong confidence level indicates the existence of a misconception (Peşman & Eryılmaz, 2010). An enhanced version of the two tiers tests are four tier tests, in which one tier measure the level of participants' confidence in the content and the addition tier measure confidence in the reason tiers.

In a lot of research, two/three/four-tier tests aiming to determine conceptual understanding in science were used. This type of test were used to determine conceptual understanding in physics (Peşman & Eryılmaz, 2010; Caleon & Subramaniam, 2010a; Kutluay, 2005; Chu et al., 2009), chemistry (Costu et al., 2007; Tan et al., 2002; Cetin-Dindar & Geban 2011; Chandrasegaran et al., 2007; Tan et al., 2002; Sreenivasulu & Subramaniam, 2013), biology (Sesli & Kara, 2012; Kılıç & Sağlam, 2009; Yen et al., 2007; Mann & Treagust, 1998; Lin, 2004) and environment (Griffard & Wandersee, 2001; Arslan et al., 2012; Cheong et al., 2015). A detailed list of studies using tier tests can be obtained in the supplementary material of Cheong et al. (2015).

With the aim to reveal common misconception of global warming (GW), greenhouse effect (GE), ozone layer depletion (OLD), and acid rain (AR) Arslan et al. (2012) develops and validate a three-tier multiple-choice diagnostic test, the atmosphere-related environmental problems diagnostic test (AREPDiT). The instrument comprised of 13 questions, each one having a standard format of three tiers: the content tier, the reasoning tier and the confidence tier. To differentiate a lack of knowledge from a misconception, a certainty response index is added as a third tier to each item. As Arslan et al. (2012) stated, the overall response possibilities to the instrument bring out five categories: Scientific knowledge, Misconception, False positives/negatives, Lucky guess, and Lack of knowledge. Results obtained in Arslan et al. (2012) study indicate that the AREPDiT is a reliable and valid instrument not only to identify pre-service teachers' misconceptions about GW, GE, OLD, and AR but also to differentiate these misconceptions from lack of knowledge. The results also indicate that a majority of the respondents demonstrated limited understandings about atmosphere related environmental problems.

Our study was conducted with three aims. The first aim of our study was to examine both construct-related validity and content validity of the AREPDiT in Bosnia and Herzegovina university student sample. The second aim of this study was to examine scientific understanding as well as misunderstanding of atmosphere-related environmental problems among B&H university student sample. Finally, the third aim of our study was to compare scientific understanding and misconceptions of the atmosphere-related environmental problems with respect to educational background.

In the line with aims, the following research questions of our study were stated: 1. What is the constructive and content validity of AREPDiT in the sample of students from the university in Bosnia and Herzegovina? 2. How do students at universities in Bosnia and Herzegovina understand environmental problems related to the atmosphere? 3. What is the understanding of these problems among students with different educational backgrounds.

METHODOLOGY

Sample

Study was conducted on 445 under-graduate students in departments of biology, geography, chemistry, mathematics, mathematics-physics and physics at the Faculty of Natural Science, Mathematics and Education in Mostar, Teachers Faculty "Džemal Bijedić" in Mostar and the Faculty of Natural Science and Mathematics in Sarajevo. In our research, 22,7% of the total examinees were male. The average age of examinees was M=20,1 (SD=0,78). Our research includes the first-year students (33,9%), the second-year students (3%), the third year students (22,7%) and the fourth year students (40,4%).

Instruments

Diagnostic test on atmospheric environmental problems (AREPDiT) Arslan et al. (2012) includes 13 tier questions on global warming, greenhouse effect, ozone layer depletion and acid rain. The first tier contains conventional multiple-choice questions that offer two to five options. The second tier offers an explanation of the answer given in the first tier. Students are given 4 to 6 possible options (one correct answer and alternative answers). The third tier determines the level of certainty in the answers given in the first two tiers. Our research determines the certainty degree via Likert 5 points scale (1 - not at all certain to 5 – absolutely certain). Topics comprised in AREPDiT are being taught to students in layered structure.

The first author of our research and an expert in climate change had analyzed contents of questions in AREPDiT from a perspective of educators that are involved in determining teaching programs that cover atmospheric environmental problems. They confirmed that topics comprised in AREPDiT are being taught to students in Bosnia and Herzegovina.

Data collection and analysis

The AREPDiT was administered to the students during the regular classes. Students completed the test in a group. The survey administration lasted approximately 25–35 min. All the statistical calculations were performed by Microsoft Office Excel and IBM SPSS Statistics v. 20.

In line with standard procedure (Pesman & Eryilmaz, 2010), based on the obtained data the following eight scores were calculated: 1) The first tiers. Only the first-tier scores are considered. The correct answer in the first tier is scored 1, incorrect 0. 2) Both tiers. The first- and second-tier scores are considered. Correct combinations are scored 1; other combinations are scored 0. 3) Total. All three tiers' scores are considered. Correct combinations of the first and second tiers with a circled "certain" and "absolutely certain" in the third tier are scored 1 point; all other combinations are scored 0 points. Total score is indicator of scientific knowledge. 4) Certainty. Only the third-tier score is considered. 5) Lack of knowledge. Combinations of responses from all three tiers are used as indicator of lack of knowledge (correct/incorrect/uncertain, incorrect/correct/uncertain, and incorrect/incorrect/uncertain). 6) Misconception first tiers (M-first tiers). 7) Misconception both tiers (M-both tiers), and 8) Misconception all tiers (M-all tiers).

RESULTS AND DISCUSSION

Validity evidence with regard to the AREPDiT

The first aim of this study was to examine both construct-related validity and content validity of the AREPDiT in Bosnia and Herzegovina university student sample. Firstly, exploratory factor analysis was conducted to identify the factor structure of the test. In addition, the correlation between score 2 and confidence levels as well as correlation between misconception score 2 and confidence levels were examined. Finally, the percentage of false positives and false negatives were calculated for content validity.

To examine construct validity of AREPDiT the exploratory factor analysis (EFA) was employed for both correct scores depending on the answers given to the three-tier questions. When factor analysis is used to test the construct validity of an instrument, it is important to take into account the measurement scale that is being used (Maydeu & D'Zurilla, 1995; Flora et al., 2003). With respect that scores are binary (correct vs. incorrect) and that distributions of items are asymmetric or with excess of kurtosis polychoric correlations, rather than the Pearson correlations is advised when carrying out this kind of analysis (Muthén & Kaplan, 1985, 1992).

The EFA was computed using FACTOR (Ferrando & Lorenzo-Seva, 2017; Lorenzo-Seva & Ferrando, 2006), a comprehensive and user-friendly program for fitting exploratory and semi-confirmatory factor analytic models (see Ferrando & Lorenzo-Seva, 2017 for conceptual view of the origins and development of FACTOR). The polychoric matrix of correlations was analyzed and principal component analysis was specified as method for components extraction. As measures of sampling adequacy of the data for conducting factor analysis, the Kaiser-Meyer-Olkin (KMO) measure and Bartlett's sphericity test were applied. The obtained results of Kaiser-Meyer-Olkin (KMO) test (0.75978) indicate fair sampling adequacy whereas value of Bartlett's statistic (χ^2=641,6, p=0,00001) indicate suitability of data for structure detection. The analysis was conducted with specifying four components due to four groups of items (GW, GE, OLD and AR). Parallel Analysis (PA) was used as a technique for determining the number of retained components. A total of 500 random polychoric correlation matrices were created using method of permutation of the raw data (Buja & Eyuboglu, 1992). Only one eigenvalue derived from the actual data (λ=5.28306) was greater compare to eigenvalues derived from the random data (λ=2.05729) indicating one significant component. This component explains 40,64% of total variance. In conclusion, the results of EFA indicate that all correct scores of three-tier questions were loaded under one component, which could be named as general scientific knowledge about atmosphere-related environmental problems.

According to Cataloglu (2002) and Pesman and Eryilmaz (2010), higher correlation between two-tier composite score and confidence tier composite score indicate construct validity of the three-tier test. If test items work properly, it could be expected that students with higher score are more confident in their answers than students with low scores (Cataloglu, 2002). Therefore, Pearson-product moment correlation coefficient between two-tier composite scores and confidence tier composite scores of the AREPDiT was calculated. Moderate positive correlation was obtained ($r= .29$, $p< .01$). It can be concluded that items of the AREPDiT work properly.

As Hestenes and Halloun (1995) claimed, the advantage of the three-tier test is the opportunity to calculate percentage of false positive and false negative to establish evidence of content validity of the test. Namely, minimization of the probability of false positive and false negative provides greater validity in multiple-choice test. Percentages of false positive and false negative were calculated by the combinations of 'correct and incorrect and certain' and 'incorrect and correct and certain' respectively. Table 1 shows the percentages of false negatives and false positives item by item.

Table 1. Percentages of false negatives and false positives

	Item													M	SD
	1	2	3	4	5	6	7	8	9	10	11	12	13		
False negative	2,7	0,7	1,1	3,4	0,7	0,2	3,2	0,2	5,6	1,1	4,8	1,4	2,5	2,1	1,7
False positive	5,0	1,4	7,0	7,2	1,4	0,5	3,4	0,2	0,7	0,2	4,8	1,1	1,4	2,6	2,5

Mean percentage of false positive (M=2,6; SD=2,5) is higher compare to mean percentage of false negative (M=2,1; SD=1,7), what is in line with Hestenes and Halloun (1995) explanation that the percentages of false positives are higher than the percentages of false negatives because reducing the probability of false positives is more difficult. The highest percentage of false negative were obtained for Item 9 (5,6%) and Item 11 (4,8%) what could be attributed to failure of some student to give sufficient attention to avoiding errors as Hestenes and Halloun (1995) suggested. When the percentages of false positives were checked, the highest percentages were obtained for greenhouse effect Item 3 (7%) and Item 4 (7,2%). Even having some misconceptions regarding greenhouse effect, students might select the correct choice in the first tier (see table 4). Taken together, it could be concluded that percentages of false negatives and false positive for all the items were smaller than 10 percentages, indicating satisfaction validity of the test.

Descriptive statistics of the AREPDiT total score

Since thorough understanding of the concepts include correct answer on content and reason tier accompanied by high confidence (Hasan, Bagayoko, & Kelley, 1999; Caleon & Subramaniam, 2010) descriptive statistics are conducted over the sum of correct responses to both first and second tiers along with being certain (correct and correct and certain). The overall descriptive statistics and the item analyses have been summarized in Table 2.

The highest score obtained in our study is 11 and the lowest is 0. The mean was M=1,77 (*SD* = 2.07) out of 13 and was very small, what indicate difficulty of the AREPDiT. The difficulty indices of the AREPDiT items indicated that the test was very difficult for students. Difficulty levels of all items except for Item 7 were below .40, with a mean of M=0.14. Skewness of the results also support this claim (Skewness=1,581, p<0,01).

Table 2. Descriptive statistics of the AREPDiT total score

Statistics	Value
Number of items	13
Number of participants	445
Mean	1,77
Median	1
Standard deviation (SD)	2,07
Minimun	0
Maximun	11
Skewness	1.581**
Kurtosis	2.805**
Cronbach alpha	0,706
Difficulty indices	
Mean	0,14
n of items < 0.10	3
n of items (0.10-0.19)	8
n of items (0.20-0.29)	1
n of items (0.30-0.39)	-
n of items (0.40-0.49)	1
Point biserial correlation	
Mean	0,480
n of items (0.30-0.39)	2
n of items (0.40-0.49)	5
n of items (0.50-0.59)	6

** p<0,01

However, item discrimination indexes indicate that items are able to distinguish between students who are knowledgeable and those who are not. As seen from Table 2, most of the point-biserial correlation coefficients are equal or above 0.30 indicating good to excellent items regarding discrimination (according to Crocker & Algina, 2008). In addition, the mean point-biserial correlation coefficient of the AREPDiT is good (r_{pb}=0.55). Item discrimination analysis indicates that the items and a test as a whole can effectively discriminate between students having high and low levels of conceptual understanding of atmosphere-related environmental problems. Finally, Croanbach's alpha reliability coefficient was found to be α= .706 which can be considered acceptable internal consistency according to criterion-referenced tests (Crocker & Algina, 2008; Kane, 1986)

Students' understanding of the atmosphere-related environmental problems

The percentages of student's correct responses according to only the content tiers, both content and reason tiers, and all three tiers as well as percentages of student's lack of knowledge and mean of confidence level are given in Table 3. The first tier shows the percentages of the student's correct responses based on the first (content) tier. Both the first and the second tiers shows the percentages of the student's correct responses to both the first (content) and the second (reason) tiers. All three tiers shows the percentages of student's correct responses to both the content and reason tiers and also were certain about their answers. Lack of knowledge is the percentages of students who were uncertain regardless of their responses. Confidence level shows the mean value of the student's certainty regardless of their response to the content and reason tier.

Table 3. Percentages of students' correct responses, lack of knowledge and mean values of confidence level

		% Correct responses			% Lack of knowledge	Confidence level
	Item	**First tiers**	**Both the first and the second tiers**	**All three tiers**		
GW	1	30,2	15,8	6,8	54,4	3,31
	2	36,0	32,5	11,1	42,0	3,31
	5	18,7	11,0	2,7	62,8	3,02
	6	24,2	23,3	14,0	45,0	3,43
	M	27,3	20,7	8,7	51,1	3,27
	SD	7,5	9,4	4,9	9,4	0,17
GE	3	77,3	50,9	20,0	35,4	3,17
	4	55,0	29,3	11,0	49,0	3,18
	M	66,2	40,1	15,5	42,2	3,18
	SD	15,8	15,3	6,4	9,6	0,01
OLD	7	68,1	61,9	40,2	22,1	3,67
	8	24,5	23,9	10,1	53,5	3,16
	9	31,8	27,9	15,7	42,5	3,42
	10	30,3	27,2	13,3	40,9	3,42
	M	38,7	35,2	19,8	39,7	3,42
	SD	19,9	17,9	13,8	13,0	0,21
AR	11	46,3	30,6	12,8	52,1	3,09
	12	13,9	9,9	6,1	64,9	3,11
	13	42,2	33,6	13,5	48,1	3,08
	M	34,1	24,7	10,8	55,1	3,09
	SD	17,6	12,9	4,1	8,8	0,02
TOTAL	M	38,4	29,1	13,6	47,1	3,26
	SD	19,0	14,5	9,1	11,3	0,19

As Table 3. illustrates, the total percentage of correct answers decreases as the number of tiers that are taken into account increases. More specifically, the mean percentage of total correct responses in content tiers was M=38,4% (SD = 19,0%), both content and reasoning tiers was M= 29,1% (SD = 14,5%,) while in all three tiers was M=13,6% (SD = 9,1%). The mean difference between the content tiers and all three tiers (M=24,7%) is noticeable. Almost quarter of students who correct answer on the content tier was unable to choose correct reason and showed confidence. The mean difference between the correct responses with regard to both and all three tiers (M=15,4) is also meaningful and correspond to students who demonstrate lack of confidence. In general, differences can be explained by lack of knowledge, lack of confidence, or misconceptions. The same pattern of decrease of percentage of correct answers as the number of tiers increase is obtain for each content area (GW, GE, OLD, and AR). As can be seen from Table 4, the students' overall understandings of each content area are low (M_{GW}=8,7%, M_{GE}=15,5, M_{OLD}=19,8, and M_{AR}=10,8). This low rate could be explained by the complex and abstract nature of the concepts of atmosphere-related environmental problems (Boyes & Stanisstreet, 1992; Cordero, 2001) as well as their educational background.

Although students' overall understandings of atmosphere-related environmental problems were very low, their confidence level were above moderate (M=3,26; SD=0,19). Overconfidence obtain in this study is in line with findings in other study in domain of psychology (Pallier et al., 2002; Renner & Renner, 2001) and educational study (Arslan et al., 2012; Caleon & Subramaniam, 2010; Taslidere, 2016).

The overall lack of knowledge mean percentage was M=47,1% (SD=11,3) indicated that almost half of the students had no understanding or were confused about their understanding of atmosphere-related environmental issues. The highest mean percentage of lack of knowledge were obtained for AR (M=55,1%; SD=8,8%) and for GW content area (M=51,1%; SD=9,4%). Somewhat lower mean percentage of lack of knowledge were obtained for GE (M=42,2%; SD=9,6%) and for OLD content area (M=39,7%; SD=13,0%).

With regard to percentages of the all three tiers score of each item, students demonstrated the higher percentage of scientific knowledge for Q7 (M=40,7%), which is about the nature of the ozone layer, and they were very certain of their answers (M=3,67). What follow is Q3 (M=20%), the question about greenhouse effect. The lowest percentages of scientific knowledge were obtained for Q5 (M=2,7%), which is about the changes in the composition of the atmosphere, and for Q12 (M=6,1%), the question about acid rain. For these two items the highest percentage of lack of knowledge were obtained (M=62,8% and M=64,9%, respectively).

In general, students' understanding of the atmosphere-related environmental problems is low. Results obtain in our study show low rate of correct responses with high certainty, what Hasan, Bagayoko, and Kelley (1999) interpreted as the indicator of misconception, an aspect which will be discussed in the next section.

Student's prevalent misconceptions

The percentages of the students holding misconceptions considering only the content tiers (M-first tiers), both content and reason tiers (M-both tiers), and all three tiers (M-all tiers) are summarized in Table 4.

Table 4. Percentage of one-tiered, two-tiered and three-tiered misconception

Misconceptions	M-first tiers	M-both tiers	M-all tiers
1. GW is caused by OLD	57,5	21,2	7,9
2. GW will cause skin cancer	23,0	22,7	8,1
3. AR is a result of GW	41,0	38,0	13,8
4. Recycling more paper is not an effective cure for GW	37,7	31,8	10,2
5. Generating electricity from renewable sources does not help to reduce GW	15,8	12,6	4,5
6. Stopping the usage of CFCs is not a cure for GW	27,5	21,6	10,9
7. GW can be reduced by setting limitations on chemical waste released into rivers	45,4	34,6	12,9
8. GW can be reduced without building nuclear power plants	21,0	19,0	9,7
9. Set a limit on pesticide usage on farmland	9,5	9,0	2,7
10. GE is not a natural phenomenon	8,3	4,5	1,1
11. Carbon dioxide (CO2) is the only gas that increases	14,2	6,3	2,2
12. GE is a totally harmful phenomenon for mankind	41,4	26,3	9,5
13. GE is a totally helpful phenomenon for mankind	1,8	0,9	0,2
14. GE has no effect on mankind	1,6	0,9	0,2
15. The ozone layer protects the Earth from AR	6,3	4,0	0,9
16. The ozone layer helps to keep the Earth's temperature stable to make it livable	25,6	16,7	6,3
17. CO2 depletes the ozone layer in the stratosphere	28,4	26,6	8,6
18. GE leads to OLD	20,7	18,3	6,1
19. Nuclear power plants affect the depletion of the ozone layer	11,5	8,3	1,1
20. Carbon monoxide (CO) causes OLD	14,9	12,0	2,9
21. OLD causes an increase in the number of floods	3,6	2,3	0,9
22. Too much sun rays enter the atmosphere by OLD	16,9	0,0	0,0
23. OLD lets the air escape from the atmosphere	2,5	0,0	0,0
24. Using public transportation reduces OLD	5,2	3,6	2,3
25. Using filters for smoke from factories and cars reduces OLD	64,5	29,7	12,8
26. OLD leads to AR	13,5	1,1	0,0
27. Methane (CH4) from landfills leads to AR	40,2	10,1	1,6
28. OLD becomes worse by AR	19,8	11,3	2,5
29. AR leads to an increase in GW	13,0	7,4	0,7
30. AR helps some plants and animals to survive	24,7	5,2	0,5
31. AR can burn everything that it comes in contact with	28,5	25,2	11,3
32. Avoiding activities that damage the ozone layer is a precaution for AR	28,2	19,4	6,5
33. CO is the main culprit of AR	29,3	25,1	7,2
M	22,5	14,4	5,0
SD	15,8	11,2	4,5

Table 4. shows that, as the number of tiers increased, the percentage of students holding misconceptions decreased. For example, 64,5% of the students hold misconception M25 (*Using filters for smoke from factories and cars reduces OLD*) according to the first tiers of the related item (item 10). However, the value decrease to 29,7% after both content and reason tiers were taken into account, indicated that 34,8% of students selected misconception alternatives in the content tiers, but not related alternatives of M25 at the reason tiers. Furthermore, when all three tiers are taken into account, 12,8% of the students select both the content and the reason tiers, and denoted confidence in terms of the confidence tiers.

The differences between mean percentage of both tiers and all three tiers indicated that16,9% of students selected the misconception alternatives in previous tiers but showed uncertainty at the confidence tiers. In line with explanation given by other authors (Arslan, Cigdemoglu, & Moseley, 2012; Hasan, Bagayoko, & Kelley, 1999; Kanli, 2014; Korur, 2015; Peşman & Eryılmaz, 2010) it seems that this 16,9% of the students had lack of knowledge rather than having corresponding misconception. Similar decreases are also noticeable with regard to all other misconception, what is expected because, as Peşman and Eryılmaz (2010) explained, student can wrongly select any specific misconception alternative at the content tier, but it is hard to find an explanation supporting that misconception at the reason tier, and to indicate confidence at the confidence tier.

Prevalent misconceptions were associated with all three tiers misconception alternatives that were selected by at least 10% of the sample, criteria proposed by Arslan, Cigdemoglu, and Moseley (2012), Caleone and Subramaniam (2010), Tan et al. (2002) and Taslidere (2016). As seen from the M-all tiers percentages in Table 4, the students hold the six prevalent misconceptions with percentages equal to or greater than 10. The prevalent misconceptions identified in this study are:

- *AR is a result of GW* (M3),
- *Recycling more paper is not an effective cure for GW* (M4),
- *Stopping the usage of CFCs is not a cure for GW* (M6),
- *GW can be reduced by setting limitations on chemical waste released into rivers* (M7),
- *Using filters for smoke from factories and cars reduces OLD* (M25),
- *AR can burn everything that it comes in contact with* (M31).

Scientific understanding and misconceptions of the atmosphere-related environmental problems with respect to educational background

The final aim of the study was to compare scientific understanding and misconceptions of the atmosphere-related environmental problems with respect to educational background. Figure 1 summarizes the mean scores for students who attend ecology classes and students who did not attend ecology classes on scientific understanding and misconception of the atmosphere-related environmental problems. The scientific understanding was calculated as a mean number of all-three tiers correct response, whereas misconception of the atmosphere-related environmental problems as mean number of M-all tiers.

Consistent with expectation, the group of students who attend ecology classes obtain higher average results of scientific understanding (M=2,16; SD=2,15) than did group of students who did not attend ecology classes (M=1,32 SD=1,86), t(443)=4,353, p<0,001. Furthermore, the group of students who attend ecology classes obtain higher average misconceptions (M=1,82; SD=1,66) than did group of students who did not attend ecology classes (M=1,46 SD=1,80), t(443)=2,188, p<0,05.

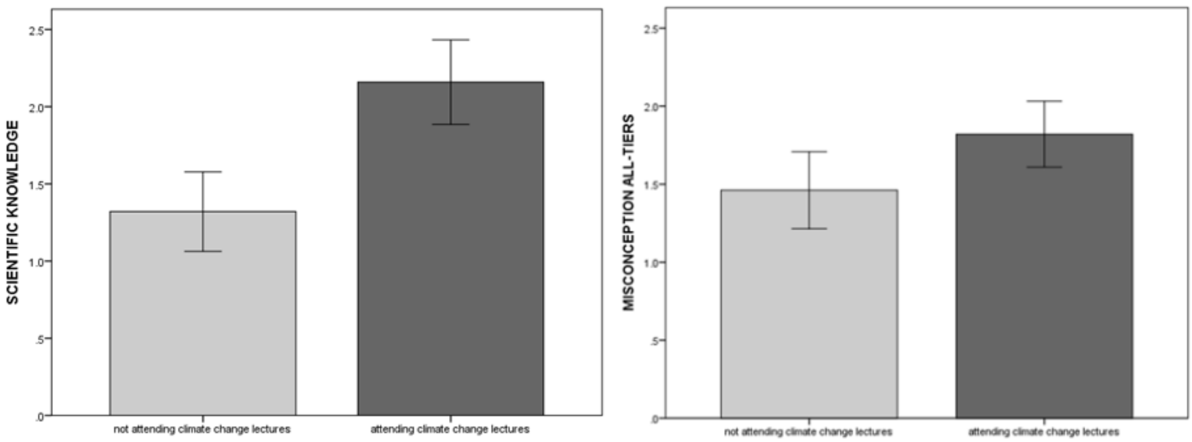

Figure 1. Means and corresponding 95% confidence intervals of the mean on the scientific understanding and misconceptions of the atmosphere-related environmental problems

CONCLUSION AND IMPLICATION

The aim of this study was threefold. Firstly, we examined both construct validity and content validity of the AREPDiT in Bosnia and Herzegovina university student sample. Further, we examined scientific understanding as well as misunderstanding of atmosphere-related environmental problem among B&H university student sample. Finally, the third aim of our study was to compare scientific understanding and misconceptions of the atmosphere-related environmental problems with respect to educational background.

For purposes of gathering evidence for construct validity, the principal component analysis was conducted on all three tiers of response. One component has been extracted which suggested that, implemented with university students in Bosnia and Herzegovina, AREPDiT measures a single construct which is general scientific knowledge about atmosphere-related environmental problems. On the other hand, in the study by Arslan, Cigdemoglu, and Moseley (2012) no reasonable factors could be obtained which was explained by loose relationship between the individual items. This once more speaks for the fact that validity is a situation-specific concept (McMilan & Schumacher, 2006) and shows that in the context of university students from Bosnia and Herzegovina it is reasonable to calculate and interpret composite scores for the whole AREPDiT scale. Further, a moderate and statistically significant correlation between correct scores and confidence scores has been found, which additionally speaks for construct validity (Cataloglu, 2002).

Content validity of AREPDiT has been checked through the analysis of false positives and false negatives. For none of the items the percentage of false positives and false negatives was above 10% which is an evidence for good content validity. Similar results were obtained by Arslan, Cigdemoglu, and Moseley (2012). However, in the study by Kahraman (2019) for two items the percentage of false positives was above 10%. We can conclude that compared to similar research, the level of content validity for our study was satisfactory.

Besides allowing us valid conclusions about students' knowledge on atmosphere-related environmental problems, it has been also found that AREPDiT scores for the Bosnia and Herzegovina university students sample may be considered as reliable.

Generally, the results of our study show that Bosnia and Herzegovina university students' knowledge about atmosphere-related problems are low but comparable to knowledge of pre-service teachers in the USA (Arslan et al., 2012; Kahraman, 2019). Similar to earlier research, most incorrect student answers resulted from lack of knowledge rather than from misconceptions. Students who attended ecology classes scored higher than their peers who had not attended these classes. However, it is even more important to note that although having higher scores on AREPDiT these same students also exhibited more misconceptions related to atmosphere-related environmental problems. This finding indicates that in the context of university education in Bosnia and Herzegovina one has to also check for possible sources of didaktikogenic misconceptions (Zajkov, Gegovska-Zajkova, & Mitrevski, 2017) related to environmental education.

As in the study conducted by Arslan et al. (2012) six misconceptions held by more than 10% of students could be identified. It is interesting to note that two of the misconceptions detected in the study by Arslan et al. (2012) were also detected in our study. One is related to the erroneous belief that acid rains are a result of global warming and the other is that chemical waste released into rivers is one of the reasons for global warming. Relating river pollution with global warming was also very common in the study by Kahraman (2019).

It is important to consider why some of the prevalent misconceptions identified in our study arise. Identifying and understanding their possible origins is critical for designing better educational materials and programs. The occurrence of these misconceptions could be related to different factors.

Students in higher education institutions rarely have the opportunity to obtain quality information related to climate change. It is often the case that curricula rely on outdated knowledge, and there is no interconnectedness between different subjects in order to address topics related to environmental issues in a clear way (Rajeev, 1997).

One of the reasons for the misconceptions related to climate change among students is the information that students receive from the media (Michail et al., 2007; Kahraman, 2020). The media most often transmit incomplete and distorted information related to environmental problems, which is justified by the task of the media to inform the public rather than educate (Dunwoody, 1992). There is also a difference in the knowledge of climate change among people who are more informed from the print media compared to those who receive news through television (Morgan & Moran, 1995). Also, most contributions related to scientific research contain a large number of factual errors leading to a misunderstanding of scientific studies (Singer & Endreny, 1993).

Fischhoff and Furby (1983) developed a research program that addresses the psychological dimensions of climate change to better understand the heuristics or mental shortcuts that people use to understand complex and uncertain threats such as climate change. This heuristic, used by both experts and lay people, often leads individuals to misjudge complex phenomena (Rajeev, 1997). Slovic et al. (1993) showed that humans are particularly vulnerable to the use of heuristics in the context of environmental risks. One example is the availability heuristic, according to which people estimate the frequency of problems according to how easily they remember certain evidence, while the other example is the representativeness heuristic, where people generalize based on limited evidence, ignoring basic statistical principles.

Although students have expressed pro-environmental views in most studies, this can lead to a combination of all adverse environmental impacts resulting in unclear knowledge related to climate change. Based on this, students can “assume” that any damage to the environment could lead to global climate change. For many students, these concepts are not yet clear enough to allow for a sophisticated consideration of cause and effect, resulting in confusion around critical issues related to climate change (Rajeev, 1997).

The effectiveness of educational messages often depends on trust in the source of the information. Rajeev (1997) states that students trust scientists the most, followed by teachers, and environmental groups and newspaper media. Government officials, family and friends had poor results on the scale with low confidence. The enormous level of trust in scientists gives both the opportunity and responsibility to the scientific community to play an active role in strengthening students’ knowledge of climate change and of the actions that students can take to prevent it (Wachholz, Artz, & Chene, 2014). Together with educators, who are also trusted, scientists can help students increase understanding and especially help them avoid the significant misconceptions found in this and other studies.

The AREPDiT can be administered to university students before and after the instruction on atmosphere-related environmental problems to gain a deeper understanding of the different views that students have on the particular topic. The results obtained by this instrument could help educators to develop the appropriate teaching-learning sequences to better address their ideas. As it is proposed by Arslan et al. (2012), the instrument could be used by educators as a tool for evaluating the effectiveness of their instruction.

This study has certain limitation. The AREPDiT is a three-tier instrument, means that it includes only one level of confidence for the answers at first and second tiers (content and reason). Therefore, it is unclear whether the confidence level is different for each tier. In the future studies a four-tier test should be applied.

REFERENCES

Anderson, B., & Wallin, A. (2000). Students' understanding of the greenhouse gas effect, the societal consequences of reducing CO2 emissions and the problem of ozone layer depletion. *Journal of Research in Science Teaching, 37*(10), 1096-1111. https://doi.org/10.1002/1098-2736

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. http://dx.doi.org/10.1080/09500693.2012.680618

Boyes, E., & Stanisstreet, M. (1992). Students' perceptions of global warming. *International Journal of Environmental Studies, 42*, 287–300. http://dx.doi.org/10.1080/00207239208710804

Buja, A., & Eyuboglu, N. (1992). Remarks on parallel analysis. *Multivariate Behavioral Research, 27*(4), 509-540. https://doi.org/10.1207/s15327906mbr2704_2

Caleon, I., & Subramaniam, R. (2010). Development and application of a three-tier diagnostic test to assess secondary students' understanding of waves. *International Journal of Science Education, 32*(7), 939–961. http://dx.doi.org/10.1080/09500690902890130

Cataloglu, E. (2002). Development and validation of an achievement test in introductory quantum mechanics: The Quantum Mechanics Visualization Instrument (QMVI). Retrieved December 24, 2019, from https://etda.libraries.psu.edu/files/final_submissions/1884

Cetin-Dindar, A., & Geban, Ö. (2011). Development of a three-tier test to assess high school students' understanding of acids and bases. *Procedia Social and Behavioral Sciences, 15*, 600-604. https://doi.org/10.1016/j.sbspro.2011.03.147

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. https://doi.org/10.1039/B7RP90006F

Cheong, I. P. A., Treagust, D., Kyeleve, I. J., & Oh, P. Y. (2010). Evaluation of students' conceptual understanding of malaria. *International Journal of Science Education, 32*(18), 2497-2519. http://dx.doi.org/10.1080/09500691003718014

Cheong, I. P. A., Johari, M., Said, H., & Treagust, D. F. (2015). What do you know about alternative energy? Development and use of a diagnostic instrument for upper secondary school science. International *Journal of Science Education, 37*(2), 210–236. http://dx.doi.org/10.1080/09500693.2014.976295

Chu, H. E., Treagust, D. F., & Chandrasegaran, A. L. (2009). A stratified study of students' understanding of basic optics concepts in different contexts using two-tier multiple-choice items. *Research in Science & Technological Education, 27*(3), 253-265. https://doi.org/10.1080/02635140903162553

Cordero, E. (2001). Misconceptions in Australian students' understanding of ozone depletion. *Melbourne Studies in Education, 41*, 85–97. https://doi.org/10.1080/17508480009556362

Coştu, B., Ayas, A., Niaz, M., Ünal, S., & Çalık, M. (2007). Facilitating conceptual change in students' understanding of boiling concept. *Journal of Science Educational Technology, 16*, 524-536. https://doi.org/10.1007/s10956-007-9079-x

Crocker, L., & Algina, J. (2008). *Introduction to classical and modern test theory*. Mason, Ohio: Cengage learning.

Dunwoody, S. (1992). The media and public perceptions of risk: How journalists frame risk stories. The Social Response to Environmental Risk, D. W. Bromley and K. Segerson, Eds., Kluwer Academic, 75-100. https://doi.org/10.1007/978-94-011-2954-1_4

Ferrando, P. J., & Lorenzo-Seva, U. (2017). Program FACTOR at 10: Origins, development and future directions. *Psicothema, 29*(2), 236-240. https://doi.org/10.7334/psicothema2016.304

Fischhoff, B., & Furby, L. (1983). Psychological dimensions of climatic change. Social Science Research and Climate Change: An Interdisciplinary Appraisal, R. S. Chen, E. Boulding, and S. H. Schneider, Eds., D. Reidel, 180-203. https://doi.org/10.1007/978-94-009-7001-4_10

Flora, B. F., Finkel, E. J., & Foshee, V. A. (2003). Higher order factor structure of a self-control test: Evidence from confirmatory factor analysis with polychoric correlations. Educational and Psychological Measurement, 63(1), 112–127. https://doi.org/10.1177/0013164402239320

Gowda, M., Fox, J., & Magelky, R. (1997). Students' understanding of climate change: Insights for scientists and educators. American *Meteorological Society, 78*(10), 2232-2240. https://doi.org/10.1175/1520-0477-78.10.2232

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. http://dx.doi.org/10.1080/09500690110038549

Groves, F. H., & Pugh, A. F. (1999). Elementary pre-service teacher perceptions of the greenhouse effect. *Journal of Science Education and Technology, 8*(1), 75–81. https://doi.org/10.1023/A:1009433705790

Groves, F. H., & Pugh, A. F. (2002). Cognitive illusions as hindrances to learning complex environmental issues. *Journal of Science Education and Technology, 11*(4), 381–390. https://doi.org/10.1023/A:1020694319071

Hasan, S., Bagayoko, D., & Kelley, E. L. (1999). Misconceptions and the certainty of response index (CRI). Physics Education, 34(5), 294–299. https://doi.org/10.1088/0031-9120/34/5/304

Hestenes, D., & Halloun, I. (1995). Interpreting the force concept inventory. *Physics Teacher, 33*, 502–506. https://doi.org/10.1119/1.2344279

Kahraman, S. (2019). Evaluating university students' understanding of atmospheric environmental issues using a three-tier diagnostic test. *International Electronic Journal of Environmental Education, 9*(1), 1-17.

Kahraman, S. (2020). Pre- and in-service teachers' understandings of atmosphere-related environmental issues: The effects of major and gender. *Erzincan University Journal of Education Faculty, 22*(1), 193-210. https://doi.org/10.17556/erziefd.615183

Kane, M. T. (1986). The role of reliability in criterion-referenced tests. Journal of Educational Measurement 23(3), 221–224. https://doi.org/10.1111/j.1745-3984.1986.tb00247.x

Kanli, U. (2014). A study on identifying the misconceptions of pre-service and in-service teachers about basic astronomy concepts. Eurasia Journal of Mathematics, *Science & Technology Education, 10*(5), 471-479. https://doi.org/10.12973/eurasia.2014.1120a

Kılıç, D., & Sağlam, H. (2009). Development of a two-tier diagnostic test to determine students' understanding of concepts in genetics. *Eurasian Journal of Educational Research, 36*, 227-244. https://doi.org/10.1016/j.sbspro.2009.01.474

Korur, F. (2015). Exploring seventh-grade students' and pre-service science teachers' misconceptions in astronomical concepts. *Eurasia Journal of Mathematics, Science & Technology Education, 11*(5). https://doi.org/10.12973/eurasia.2015.1373a

Kutluay, Y. (2005). Diagnosis of eleventh grade students' misconceptions about geometric optic by a three-tier test. Unpublished master thesis, Middle East Technical University, Ankara.

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*, 175-199. https://doi.org/10.1007/s10763-004-6484-y

Lorenzo-Seva, U., & Ferrando, P. J. (2006). FACTOR: A computer program to fit the exploratory factor analysis model. Behavior Research Methods, 38(1), 88-91. https://doi.org/10.3758/BF03192753

Mann, M., & Treagust, D. (1998). A pencil and paper instrument to diagnose students' conceptions of breathing, gas exchange and respiration. Australian Science Teachers Journal, 44, 55-60.

Maydeu, A., & D'Zurilla, T. J. (1995). A factor analysis of the social problem-solving inventory using polychoric correlations. *European Journal of Psychological Assessment, 11*(2), 98–107. https://doi.org/10.1027/1015-5759.11.2.98

McMillan, J. H., & Schumacher, S. (2006). *Research in education: Evidence-based inquiry (6th ed.)*. Boston, MA: Allyn and Bacon.

Michail, S., Stamou, A. G., & Stamou, G. P. (2007). Greek primary school teachers understanding of current environmental issues: An exploration of their environmental knowledge and images of nature. *Science Education, 91*(2), 244–259. https://doi.org/10.1002/sce.20185

Morgan, M. D., & Moran, J. M. (1995). Understanding the greenhouse effect and the ozone shield: An index of scientific literacy among university students. *The Bulletin of the American Meteorological Society, 76*, 1185-1190. https://doi.org/10.1175/1520-0477-76.7.1185

Muthén, B., & Kaplan D. (1985). A comparison of some methodologies for the factor analysis of non-normal Likert variables. *British Journal of Mathematical and Statistical Psychology, 38*, 171-189. https://doi.org/10.1111/j.2044-8317.1985.tb00832.x

Muthén, B., & Kaplan D. (1992). A comparison of some methodologies for the factor analysis of non-normal Likert variables: A note on the size of the model. *British Journal of Mathematical and Statistical Psychology, 45*, 19-30. https://doi.org/10.1111/j.2044-8317.1992.tb00975.x

Pallier, G., Wilkinson, R., Danthiir, V., Kleitman, S., Knezevic, G., Stankov, L., & Roberts, R. D. (2002). The role of individual differences in the accuracy of confidence judgments. The *Journal of general psychology, 129*(3), 257–299. http://dx.doi.org/10.1080/00221300209602099

Papadimitriou, V. (2004). Prospective primary teacher's understanding of climate change, greenhouse effect, and ozone layer depletion. *Journal of Science Education and Technology, 13*(2), 229-307. https://doi.org/10.1023/B:JOST.0000031268.72848.6d

Paterson, M. (1996). *Global warming and global politics*. London: Routledge.

Pekel, O. F., & Ozay, E. (2005). Turkish high school students' perceptions of ozone layer depletion. *Applied Environmental Education and Communication, 4*, 115–123. http://dx.doi.org/10.1080/15330150590934598

Peşman, H., & Eryılmaz, A. (2010). Development of a three-tier test to assess misconceptions about simple electric circuits. *The Journal of educational research, 103*(3), 208-222. https://doi.org/10.1080/00220670903383002

Pruneau, D., Liboiron, L., Vrain, E., Gravel, H., Bourque, W., & Langis, J. (2001). People's ideas about climate change: A source of inspiration for the creation of educational programs. *Canadian Journal of Environmental Education, 6,* 121-138.

Rajeev Gowda, M. V., Fox, J. C., & Magelky, R. D. (1997). Students' understanding of climate change: Insights for scientists and educators. *Bulletin of the American Meteorological Society, 78*(10), 2232-2240. https://doi.org/10.1175/1520-0477-78.10.2232

Renner, C. H., & Renner, M. J. (2001). But I thought I knew that: Using confidence estimation as a debiasing technique to improve classroom performance. *Applied Cognitive Psychology, 15*(1), 23–32. https://doi.org/10.1002/1099-0720(200101/02)15:1<23::AID-ACP681>3.0.CO;2-J

Rye, J. A., Rubba, P. A., & Wiesenmayer, R. L. (1997). An investigation of middle school students' alternative conceptions of global warming. *International Journal of Science Education, 19*(5), 527–551. http://dx.doi.org/10.1080/0950069970190503

Salinger, M. (2005). Climate variability and change: past, present, and future – an overview. *Climate Change, 70*(1), 9-29. https://doi.org/10.1007/s10584-005-5936-x

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. http://dx.doi.org/10.1080/00219266.2012.688849

Singer, E., & Endreny, P. M. (1993). *Reporting on Risk: How the Mass Media Portrays Accidents, Diseases, Disasters, and Other Hazards*. Russell Sage Foundation.

Slovic, P. (1993). Perceived risk, trust, and democracy. *Risk Analysis,13*, 675-682. https://doi.org/10.1111/j.1539-6924.1993.tb01329.x

Sreenivasulu, B., & Subramaniam, R. (2013). University students' understanding of chemical thermodynamics. International *Journal of Science Education, 35*(4), 601-635. http://dx.doi.org/10.1080/09500693.2012.683460

Summers, M., Kruger, C., Childs, A., & Mant, G. (2000). Primary school teachers' understanding of environmental issues: An interview study. *Environmental Education Research, 6*(4), 293–312. http://dx.doi.org/10.1080/713664700

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: The Official Journal of the National Association for Research in Science Teaching, 39*(4), 283-301. https://doi.org/10.1002/tea.10023

Taşlıdere, E. (2016). Development and use of a three-tier diagnostic test to assess high school students' misconceptions about the photoelectric effect. *Research in Science &*

*Technological Education, 34(*2), 164-186. http://dx.doi.org/10.1080/02635143.2015.1124409

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

United Nations Intergovernmental Panel on Climate Change (IPCC) (2007). Climate Change 2007: Synthesis Report. Cambridge University Press, Cambridge, MA.

Wachholz, S., Artz, N., & Chene, D. (2014). Warming to the idea: University students' knowledge and attitudes about climate change. *International Journal of Sustainability in Higher Education, 15(*2), 128-141. http://dx.doi.org/10.1108/IJSHE-03-2012-0025

Yen, C., Yao, T., & Mintzes, J. J. (2007). Taiwanese students' alternative conceptions of animal biodiversity. *International Journal of Science Education, 29*(4), 535-553. http://dx.doi.org/10.1080/09500690601073418

Zajkov, O., Gegovska-Zajkova, S., & Mitrevski, B. (2017). Textbook-caused misconceptions, inconsistencies, and experimental safety risks of a grade 8 physics textbook. *International Journal of Science and Mathematics Education 15*, 837–852. https://doi.org/10.1007/s10763-016-9715-0