JOURNAL OF CHEMICAL EDUCATION

Article

pubs.acs.org/jchemeduc

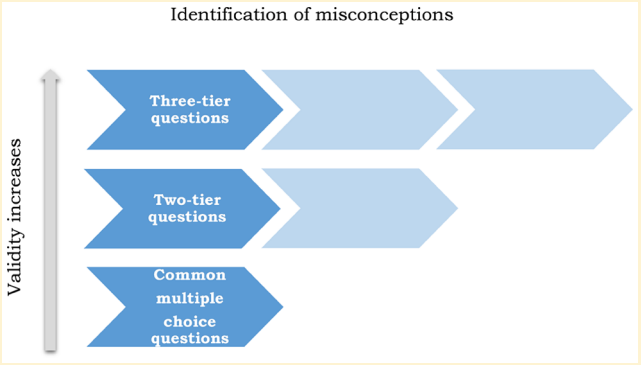
Development of a Three-Tier Test as a Valid Diagnostic Tool for Identification of Misconceptions Related to Carbohydrates

Dušica D. Milenković,* Tamara N. Hrin, Mirjana D. Segedinac, and Saša Horvat

Department of Chemistry, Biochemistry, and Environmental Protection, Faculty of Sciences, University of Novi Sad, Novi Sad 21000, Republic of Serbia

Supporting Information

ABSTRACT: This study describes the development and application of a three-tier test as a valid and reliable tool in diagnosing students' misconceptions regarding some basic concepts about carbohydrates. The test was administrated to students of the Pharmacy Department at the University of Bijeljina (Serb Republic). The results denoted construct and content validity of the instrument and its high reliability. Considering the results entirely, it was determined that the presence of a third tier in a task significantly affected the number of identified misconceptions and the assessment of scientific knowledge. These findings suggest that certain incorrect responses should be attributed to the lack of knowledge rather than to misconceptions, and likewise, some correct responses should be attributed to the lucky guess rather than to scientific knowledge, which implies a higher credibility of the results obtained. In this study, several misconceptions about carbohydrates were revealed. As there are only a few studies examining students' misconceptions in chemistry using three-tier tests available in literature, future research should focus on the further implementation of three-tier tests as a powerful tool for simple detection of misconceptions.



KEYWORDS: *Upper-Division Undergraduate, Biochemistry, Organic Chemistry, Misconceptions/Discrepant Events, Testing/Assessment, Carbohydrates*

FEATURE: Chemical Education Research

INTRODUCTION

According to the constructivist theory of learning, people are able to create and develop cognitive models to explain everyday life phenomena. By that, to generate their own models they rely on the attitudes, skills, and experiences, as new information starts to make sense only if it can be incorporated into existing schematic model of the individual. According to information obtained from teaching practice, students' constructions are quite idiosyncratic. Namely, after a new concept has been introduced, each student is likely to form his own version of the same, which will to some extent differ from others.[1] These concepts are often different from those that the teacher wants to present, that is, from scientifically accepted concepts.

During the formation of their own constructions, students invest a significant amount of mental effort, and therefore, once created and adopted, a misconception is hard to eliminate and to replace with a proper, scientifically accepted concept. Besides that, persistency of misconceptions is additionally stimulated by their simplicity and intuitive clarity. On the contrary, chemistry is a subject that is, to a large degree, based on abstract concepts and therefore difficult to understand and learn, especially when students are put in a position to believe in something they cannot register by senses.[2] Since misconceptions significantly hinder the learning process, researchers and educators in the field of chemical education have extensively worked, during the last several decades, to gather information on students' misconceptions, aiming to eliminate them.

THEORETICAL FRAMEWORK

The procedures for identifying misconceptions have been changing and improving over time. In the beginning, research studies commonly included interviews as they provided in-depth information about students' cognitive models.[3−5] However, this form of diagnosing misconceptions has proven to be very time-consuming and thus uneconomical. Namely, their preparation and administration require a lot of time, which significantly reduces the number of potential study subjects. Therefore, perceived misconceptions could be attributed to individuals, and do not necessarily reflect the state of the group. Another important issue related to interviewing is an interviewer. A successful interview implies having a skilled interviewer who does not possess any cognitive biases that could hinder the

Received: April 7, 2016
Revised: June 16, 2016

DOI: 10.1021/acs.jchemed.6b00261
J. Chem. Educ. XXXX, XXX, XXX−XXX

research. Finally, the analysis of the results itself can sometimes be difficult and abstruse.[6]

Aiming to save the time, researchers have begun to use written forms of knowledge evaluation. Tests with open-ended questions were first in use.[7,8] They were considered convenient as students formulated answers by themselves, which enabled researchers to determine whether the examined student possessed misconceptions or not. However, over time, this form of misconception identification proved inefficient as well, due to the available teaching time, only a small number of open-ended questions could be done. Furthermore, it was found that students were reluctant to write full sentences and give detailed answers during testing.[9] For this reason, multiple-choice questions became quite popular among researchers and widely implemented in research designs.[10−12] It is important to note that during the construction of such tests, the quality of the distractors should be taken into account. Namely, distractors have to represent some potential misconceptions in order to be appealing to students.

These tests are very efficient in terms of time, because they can cover a large number of subjects and include a large number of tasks. However, this format is quite often criticized due to the high probability of guessing the correct answer. For this reason, researchers have developed a new form of testing, that is, tests consisting of two-tier tasks.

Through the pioneering work of Treagust[13,14] and later[15−18] two-tier tests became the prevalent way of identifying misconceptions among students. Two-tier tasks consist of two parts. The first part contains the content problem, while the second part contains a reasonable explanation of the problem presented in the first tier of the task.[19] Unlike a common multiple choice task with four options, only one of which is correct, where the probability of guessing the correct answer is 25%, in two-tier tasks probability of guessing is significantly reduced and equals approximately 6%.

The use of two-tier test allows teachers and researchers not only to understand students' misconceptions, but also to explore the reasoning behind them. In addition to this, these tests facilitate the examination and evaluation of misconceptions on a large number of subjects in an efficient and simple manner, since they are suitable for administration and the time needed for their realization minimally encroaches on the available teaching time.[19]

Based on the results achieved on the two-tier test it is possible to determine whether students possess some misconceptions and to determine the level of conceptual understanding. According to Gilbert[20] satisfactory conceptual understanding exists if at least 75% of the students choose the correct answer (for tasks with four offered answers). If the percentage is lower (50−74%), then it can be said that students achieved a roughly adequate performance. The percentage of choosing the correct answer between 25 and 49% indicates an inadequate performance, while the percentage of choosing the correct answer less than 25% indicates quite an inadequate performance. On the other hand, according to this author, based on the percentage of choosing distractors it is possible to determine the existence of misconceptions among students. Thus, a distractor-choosing frequency greater than 20% can serve as a reliable indicator of the existence of misconception in the examined group of students.

Although two-tier tests have numerous advantages over ordinary multiple-choice tests, there are certain limitations that should be mentioned. Results obtained through two-tier tests do not differ between misconceptions and lack of knowledge as well as between understanding and lucky guesses. For that reason, researchers were encouraged to develop more complex form of multiple-tier test which in addition to content and reason tiers contains an additional tier, the so-called "confidence" tier. The application of confidence judgments in research designs is widely documented in literature,[21−24] which supports the validity of this method.

Introducing and considering the third tier in tasks provides valuable information about students' self-confidence. Namely, only those students who provide the correct answers in both tiers, and indicate that they are sure of their answers, understand the content of the task. Otherwise, if a student is not sure of the answers provided for the first and second tier, it can be concluded that the correct answers are the result of guessing. On the other hand, only those students who give incorrect responses in both tiers or in one of the tiers, and state that they are sure of their answers, possess a misconception, otherwise incorrect responses are likely to be the result of a lack of knowledge. This greatly reduces the number of students who are identified as holding misconceptions, as well as students who are identified as possessing scientific knowledge, which significantly increases the validity of the study results.

It is interesting (see ref 9) that most research studies which used a three-tier test of knowledge as an instrument for the identification of misconceptions were carried out on the contents of physics,[25−33] while in the case of chemistry there are only two such studies.[34,35] Additionally, reviewing the relevant literature it can be observed that a number of studies have published findings which describe general and inorganic chemistry misconceptions, while investigations of organic chemistry misconceptions have been somewhat neglected.[36−38] According to our knowledge, there are very few studies regarding carbohydrates.[39,40] Hence, the current study aims to enrich both current literature on organic chemistry misconceptions and the application of three-tier test on chemistry contents.

METHODOLOGY

Research Objective

The main objective of this research was to develop a three-tier test of knowledge as a valid diagnostic tool for the identification of misconceptions in the topic carbohydrates.

Participants

Testing was performed at the Pharmacy Department, University of Bijeljina (Bijeljina, Serb Republic). It is a small private institution that was selected for this research primarily due to its size (number of students). Namely, this study included all second and third year students (N = 42) who had accepted to voluntarily participate in the study. Population by gender was as follows: 19% males and 81% females. Since there is a lack of information on students' misconceptions about carbohydrates in current literature, to provide valid task distractors, the authors of this paper considered it convenient to engage an organic chemistry teaching assistant who is familiar with the students' knowledge as well as their knowledge gaps in instrument construction. To make this possible, a small size university was chosen, as only in that way the teaching assistant could have a detailed insight into the students' knowledge. More detailed explanation is provided in the Research Instrument Section.

According to the curriculum, the teaching topic Carbohydrates is studied within the course Organic Chemistry II. This course is taught in the third semester (15 weeks), with 30 h of lectures and 30 h of laboratory practice. During the semester,

DOI: 10.1021/acs.jchemed.6b00261
J. Chem. Educ. XXXX, XXX, XXX−XXX

students collect pre-exam points comprised of presence points, a seminar report, and three compulsory colloquiums, which altogether amounts to 50 points. At the end of semester, the students have an oral exam worth 50 points.

Research Instrument

For the purpose of this research, a carbohydrate diagnostic test of knowledge (CDTK) was constructed. A teaching assistant for the organic chemistry courses, at the university where the research was conducted, actively participated in construction of the CDTK. Namely, she was versed in the contents taught within this course and familiar with level of the students' progress. What is more important, she taught 2 h of lab sessions each week over a 15-week semester within the course Organic Chemistry I, and 2 h of lab sessions each week within the course Organic Chemistry II, which included continuous weekly discourses, thus giving her insight into the issues and potential misconceptions that students could have. Based on her observations and the results of the mandatory tests during the semester, the task distractors were designed.

The CDTK contained 14 tasks, each of which consisted of three tiers. The first tier was a multiple-choice content question that consist of 4 options, only one of each was correct, while the remaining three were distractors. The second tier offered 4 answers as well, only one of each was correct, while the remaining three distractors represented logical explanations of the distractors given in the first tier. Finally, the third tier of the question represented the confidence tier, which serves as a determinant of the students' confidence in their answers provided for the first and second tier. Further in the text (Box 1), we provide an example item (the complete test is given in Supporting Information Section).

Box 1. Example of Test Item

Circle the letter of the correct answer. Which of the following monosaccharides in the reaction with sodium borohydride yields the same polyhydric alcohol as glucose?

a. Galactose
b. Fructose
c. Ribose
d. Glyceraldehyde

The reason for your answer is

a. Molecules of the selected monosaccharide contain one carbon atom less than glucose molecule, and in the reaction with sodium borohydride the number of carbon atoms increases by one
b. Molecules of the selected monosaccharide contain the same number of carbon atoms and aldehyde group as molecules of glucose
c. Molecules of the selected monosaccharide contain keto group on C2 which is transformed into a hydroxyl group in the reaction with sodium borohydride
d. Molecules of the selected monosaccharide contain three carbon atoms and in reaction with sodium borohydride, the number of carbon atoms doubles

Are you sure of your answers?

a. Yes
b. No

To determine the validity of the CDTK, standard procedures that involve calculating the pretest and post-test quality assurance parameters, were used. Within the pretest assurance parameters an expert team, comprised of two university professors and three teaching assistants in the field of chemistry education, considered readability, meaningfulness of requirements, and compliance with the content of the curriculum and concluded that the CDTK is a valid instrument. Within the post-test parameters: (i) reliability expressed by Cronbach α coefficient for the entire test and separately for each tier; (ii) indices of task difficulty; (iii) point biserial correlation coefficients; and (iv) correlation between both tiers score and certainty (third tier scores) were calculated. The main results of the analysis are presented in the Results and Discussion Section.

Data Collection

Testing was conducted in September 2015. The students had 45 min to administer the test. All tests were carefully reviewed and the results entered into Microsoft Office Excel. All the statistical calculations were performed by Microsoft Office Excel and IBM SPSS Statistics v. 20.

Based on the obtained data, the following parameters were calculated:

1. First-tier score [FT] (Only the first-tier scores are considered. The correct answer in the first tier is scored 1, incorrect 0.)
2. Second-tier score [ST] (Only the second-tier scores are considered. The correct answer in the second tier is scored 1, incorrect 0.)
3. Third-tier score [TT] is a certainty score. (Only the third-tier score is considered. The answer "yes" is scored 1, the answer "no" is scored 0.)
4. Both-tiers score [BT] (The first- and second-tier scores are considered. Correct combinations are scored 1; other combinations are scored 0.)
5. All-tiers score [AT] (All three tiers' scores are considered. Correct combinations of the first and second tiers with a circled "yes" in the third tier are scored 1 point; all other combinations are scored 0 points.)

Based on the calculated scores, possible combinations of answers were defined. If a student answers both the first and the second tier correctly, and states YES in the third tier, then it is possible to conclude that the student possesses a scientific knowledge of the examined concept. If a student answers the first tier correct, the second incorrect, and circles YES in third, then it can be assumed that the student possesses a misconception, and such students are called false positives. Likewise, when a student gives an incorrect response in the first tier, but has the correct reasoning in the second tier and circles YES in the third, the student probably possesses a misconception and such students are called false negatives. We can be certain that a student possesses a misconception if both the first and second tiers are answered incorrectly and YES is circled in the third tier. Another possibility is that a student answers both the first and second tiers correctly, but circles NO in the third. That is, most likely, the result of a lucky guess or lack of confidence. All other combinations (correct−incorrect−no; incorrect−correct−no; and incorrect−incorrect−no) indicate a lack of knowledge.

RESULTS AND DISCUSSION

Post-Test Assurance Parameters

As previously mentioned, before the main analysis, an analysis of the reliability of the CDTK was performed. The first step in

DOI: 10.1021/acs.jchemed.6b00261
J. Chem. Educ. XXXX, XXX, XXX−XXX

the assessment of CDTK reliability included the calculation of the value of Cronbach α coefficient for achievements accomplished for the FT, BT, and AT. All three obtained values are quite high, for the FT 0.735, for the BT slightly higher 0.797, while the AT coefficient of reliability was the highest and equals 0.887. The next step was related to the calculation of tasks' difficulty index, considering the FT, BT, and AT scores. This analysis showed that the mean item difficulty for the FT is very high (0.72), for the BT it is lower (0.67), while for the AT it is, expectedly, the lowest (0.47). Regarding the AT scores, item difficulty analysis has shown that there are two tasks with a difficulty index less than 0.30, which characterizes them as difficult tasks. All the remaining tasks are in the range 0.30−0.80 and those are the tasks of moderate difficulty.

The second part of the item analysis involved calculations of point biserial correlation coefficients for the AT scores. The results of this analysis showed that there are only two tasks with point biserial coefficients between 0.20 and 0.30 (classifying them as appropriate), while the remaining 12 tasks have a very good value of this coefficient (over 0.30). The average value of this coefficient for all tasks is very high and equals 0.55. All the above indicators suggest a good quality of the constructed tasks, that is, of the CDTK.

For the validation of the applied procedure for analysis of misconceptions, a common method proposed by Cataloglu (according to ref 41) was used. Namely, in his research, the above-mentioned researcher found that those students who accomplished higher achievements were more certain of their answers, while those who achieved lower results were accordingly, less confident in their answers. Thereby, the basis of the method for providing the evidence of construct validity of the CDTK is the correlation between the BT and TT scores by the students. In this research, strong and positive correlation between the BT and TT scores (r = 0.75; n = 42; p = 0.00) was computed. Correlation between the BT and TT scores is shown graphically in Figure 1. The graph shows that with

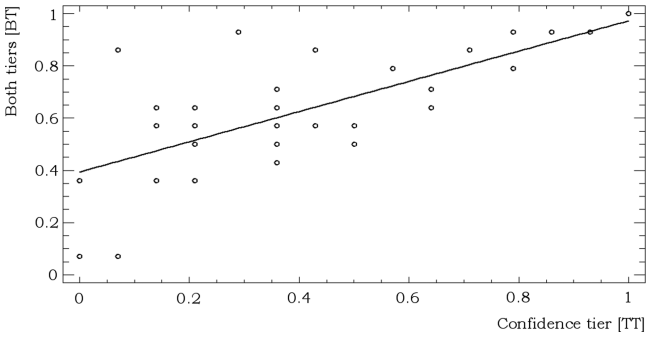


Figure 1. Correlation between both-tiers (BT) and third-tier (TT) scores.

increasing confidence level, there is an increase in the number of correct answers in both tiers. However, it is important to note that there were those students who were unsure of their answers, even though they achieved very high scores on the test. This implies a lack of confidence.

According to the relevant literature in the field of three-tier test evaluation,[27,41] content validity is additionally affirmed by calculating the percentage of false negatives and false positives. It is recommended that the percentage of false negatives should not exceed 10%. In this study the obtained percentage of false negatives is 5.8% and of false positive is 5.9%, which are in the range of recommended values. Based on the results obtained, it can be concluded that the CDTK represent a valid instrument and can be used for further identification of misconceptions.

Main Statistics and Analysis of Misconceptions

Table 1 summarizes the overall statistics based on the AT scores. The values of standardized skewness and standardized kurtosis

Table 1. Descriptive Statistics for All-Tier Scores

parameter	value
number	42
average score	6.64
standard deviation	4.11
minimum score	0.00
maximum score	14.00
range	14.00
standard skewness	0.54
standard kurtosis	−1.49

are within the values that correspond to normal distribution. The average value of the achievements for AT is satisfactory (47.45%) and it is slightly less than half of the maximum possible performance. Based on the data of the maximum achieved results, we have briefly gone through the data matrix and noted that there was only one student who gave all the correct answers, and who was sure of all the answers on the test.

Analysis of the correct answers for all tasks and for all subjects involved the comparison of student achievements realized in the FT, then BT and AT. The results of this analysis are presented in Table 2.

Table 2. Results of the Analysis of Correct Answers by Tiers

	score averages by test-tier, N = 42		
question	first tier	both tiers	all tiers
1	69.05	66.67	38.10
2	71.43	69.05	61.90
3	83.33	83.33	64.29
4	83.33	61.90	19.05
5	66.67	66.67	38.10
6	97.62	78.57	54.76
7	80.95	73.81	71.43
8	61.90	59.52	57.14
9	78.57	73.81	30.95
10	9.52	2.38	2.38
11	78.57	78.57	73.81
12	66.67	54.76	35.71
13	76.19	76.19	64.29
14	92.86	90.48	52.38
mean	72.62	66.84	47.45

If we compare the average values of achievements shown in Table 2 it can be easily observed that achievement gradually decreases as the number of tiers increases. These results are, in a way, expected as they are in a full agreement with previous research (see ref 9). The difference between the scores obtained for the FT and BT is approximately 6% and could be attributed to the existence of false positives. On the other hand, the difference in the mean value achieved for BT and AT is slightly over 19% and could be explained either by a lucky guess or lack of confidence. The significant difference in the scores achieved for the FT and AT (over 25%) can be attributed to a lucky guess, lack of knowledge, and misconceptions.

DOI: 10.1021/acs.jchemed.6b00261
J. Chem. Educ. XXXX, XXX, XXX−XXX

These results confirmed the fact that students are solving FT issues quite easily, while BT issues are much harder. Furthermore, it is easy to notice that there are students who were not sure of their answers, even when they provided correct answers to the BT. Accordingly, it is evident that common multiple choice questions and even two-tier tests have certain shortcomings as evaluation tools, that is, these tests overestimate students' understanding and by introducing a third tier, more credible results can be obtained.

In addition to the analysis of achievements by tiers, the quality of students' knowledge was tested. The percentage of students who possess scientific knowledge (SK), lack of knowledge (LK), misconceptions (M), and the percentage of students who provided the correct answers by a lucky guess (LG) were calculated. These results are summarized in Table 3.

Table 3. Analysis of Students' Conceptual Knowledge

	student test response categorizations,[a] % (N = 42)			
question	SK	LG	LK	M
1	38.10	28.57	26.19	7.14
2	61.90	7.14	19.06	11.90
3	64.29	19.05	14.28	2.38
4	19.05	42.86	38.09	0.00
5	38.10	28.57	30.95	2.38
6	54.76	23.81	16.67	4.76
7	71.43	2.38	21.43	4.76
8	57.14	4.76	23.81	14.29
9	30.95	42.86	21.43	4.76
10	2.38	0.00	88.10	9.52
11	73.81	4.76	9.53	11.90
12	35.71	19.05	33.34	11.90
13	64.29	11.90	9.52	14.29
14	52.38	38.10	7.14	2.38
mean	47.45	19.56	25.68	7.31

[a]Category abbreviations indicate that students had: SK, scientific knowledge; LG, a lucky guess; LK, a lack of knowledge; M, misconceptions.

Considering the results obtained for scientific knowledge, it can be observed that the values range from very low (Q10) to very high (Q11). However, the average value is moderate (47.28) and shows that there is a relatively high percentage of students who successfully solved both tiers of the task and who were certain of their answers. Without considering the third tier, this percentage would be even higher, because all students who provided the correct answer by a lucky guess would be considered as students who understand the content.

Hence, three-tier tests provide more accurate data, as only those students who are certain of their answers can be said to possess the scientific knowledge of the examined content.

Since the examination of misconceptions was the main task of this study, it was performed in detail by comparative analysis of the results based on: (i) the FT only (as if the test were in a form of conventional multiple-choice test); (ii) BT (as if the test were in a form of two-tier test); and (iii) AT (three-tier test). The results of this analysis are graphically presented (Figure 2). This figure clearly shows that the percentage of misconceptions decreases with increasing number of tiers in the task. According to this notation, it can be concluded that the application of three-tier tests allows more accurate identification of misconceptions as this method enables researchers to distinguish between misconceptions and lack of knowledge. These results

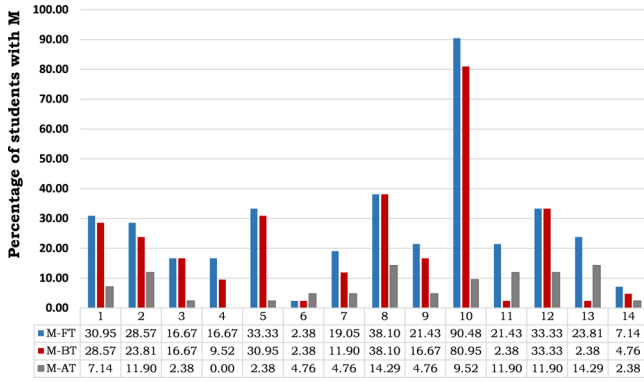


Figure 2. Analysis of misconceptions by tiers.

replicate previous findings (see ref 9) and recommend three-tier tests as superior tools for diagnosing misconceptions in comparison to widely used two-tier and conventional multiple-choice question tests.

According to Figure 2, it can be observed that there are tasks with a high percentage of incorrect answers in the FT and BT (e.g., Q1, Q5, Q8, Q10, and Q12). However, this number is significantly reduced when a third tier is taken into consideration, as the results of the third tier provide information about the cause of the incorrect answers, that is, whether they are a result of a misconception or lack of knowledge. Hereinafter we will analyze the tasks in which some interesting misconceptions have been recorded.

In Q1 (see the Supporting Information) the students were required to conclude which claim relates to the cyclic form of D-(−)-ribose. To solve this task, the student needs to know the structure of molecules of D-(−)-ribose, cyclization of molecules of carbohydrates, relative configuration, the concept of reducing and nonreducing carbohydrates, chirality, and optical activity. In the first tier of this question, the students chose distractor d (has 3 asymmetric centers) in a large percentage, while in the second tier they chose its pair distractor a (cyclic molecules of D-(−)-ribose contain 5 C atoms the first and last of which are not asymmetric). These results showed that the students reflected their knowledge about acyclic form the D-(−)-ribose to a cyclic form, thus believing that cyclization of molecule of D-(−)-ribose, which consists of 5 carbon atoms, the first and last of which are not asymmetric, results in forming cyclic molecule which also has three asymmetric centers. If we consider the third tier, we can see that only 7.14% of students are certain of their answers and only for these students, we can be certain that they possess a previously described misconception.

In Q5, students were required to respond indicating which monosaccharide in reaction with sodium borohydride will yield the same polyhydric alcohol as glucose. In order to successfully solve this task, students had to know that sodium borohydride is a reducing agent and to know the orientation of the hydroxyl groups in a given molecule. The majority of respondents wrongly concluded that it was a galactose, with explanation that glucose and galactose molecules contain the same number of carbon atoms and an aldehyde group ("the same functional group").

Q8 referred to the determination of mass percentage of carbon in molecules of different monosaccharides (vide supra). To solve this task, the respondent should have known that the general formula of monosaccharide molecule is $C_nH_{2n}O_n$, and that the mass percentage of carbon in a molecule of each monosaccharide is the same, regardless of the number of carbon

DOI: 10.1021/acs.jchemed.6b00261
J. Chem. Educ. XXXX, XXX, XXX−XXX

atoms in its molecule. However, most respondents circled mannose as a correct answer, as molecules of mannose contain the highest number of carbon atoms. When all three tiers are taken into account, Q8 is a task with the largest percentage of respondents who possess misconception.

In the text of Q10 the students were given the fact that fructose is the sweetest natural sugar, and were expected to use that fact and compare the sweetness of honey and white sugar. The lowest percentage of correct answers in both the FT and BT was recorded in this task, and the most frequently chosen distractor was that honey and white sugar are equally as sweet as sucrose and invert sugar building blocks are composed of the same units (glucose and fructose molecules). Thus, students showed that they did not consider the fact that the molecules of white sugar, unlike invert sugar, consist of chemically bonded glucose and fructose molecules. However, the number of misconceptions significantly reduces if the third tier is taken into consideration, because a large percentage of misconceptions estimated by the BT scores can be attributed to the lack of knowledge.

In Q12 respondents were supposed to answer the question what is determined by the orientation of −H and −OH groups around the carbon at a position 5 in the molecules of aldohexoses. Although a high percentage of the respondents gave the correct answers to this question, there were those who chose other options. The most frequently selected distractor is the distractor b, optical activity, with the following explanation: aldohexoses in which the orientation of −OH group in the position 5 is the same as orientation of −OH group attached to the asymmetric carbon atom of D-glyceryl aldehyde, rotates the plane-polarized light to the right, while those aldohexoses in which orientation of −OH group in the position 5 is the same as orientation of the −OH group attached to the asymmetric carbon atom of L-glyceryl aldehyde, rotates the plane-polarized light to the left. It can be assumed that the students have associated D and L marks with the rotation direction of polarized light (in Serbian, *Desno* means "right", and *Levo* means "left"). Considering the third tier, it can be concluded that the respondents were largely confident of their answers.

CONCLUSION

This study is focused on the creation of three-tier test (CDTK) and determination of its quality for diagnosing students' misconceptions about carbohydrates. The results showed that the CDTK has moderate difficulty and that it represents a valid and reliable instrument for identifying misconceptions and students' understanding with high level of certainty. Findings of this study are consistent with previous research, which showed that three-tier instruments are more appropriate than two-tier and common multiple-choice tests, as they identify misconceptions with a higher reliability, distinguishing them from the lack of knowledge. Additionally, three-tier instruments provide more reliable feedback on students' understanding, distinguishing this from lucky guesses. This is made possible by applying the third tier that provides data about students' self-confidence.

The significance of this research is multifaceted. Namely, although the research on misconceptions has been attracting the attention of researchers for several decades, it is worth mentioning that there is almost no research on misconceptions regarding carbohydrates. Therefore, this research, which identified several misconceptions, contributes to the enrichment of literature in this area. Another important contribution of this research is reflected in the application of the three-tier test as a tool for assessing misconceptions and conceptual understanding of students, which is underutilized in the contents of chemistry.

However, this study has certain limitations that should be mentioned. They relate to the lack of three-tier instruments. In the first place, that is the third tier, the level of confidence, which applies to both the first and second tier. If the confidence tier had been split into two parts, that is, if students had evaluated separately how confident they were in their answers to the first and second tier, it would have been possible to obtain even more precise data. That could provide a finer structuring of the results and avoid the possibility of underestimating the proportion of the lack of knowledge and overestimating the students' scores.

Regarding the implications, further research should begin with an increment of sample size and widen the research to other content areas (primarily organic chemistry, which is insufficiently researched). Additionally, in accordance with the previously mentioned limitation, the third tier should be separated in two parts in order to provide more precise results.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available on the ACS Publications website at DOI: 10.1021/acs.jchemed.6b00261.

Research instrument (PDF, DOC)

AUTHOR INFORMATION

Corresponding Author

*E-mail: dusica.milenkovic@dh.uns.ac.rs.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was funded by the Ministry of Education, Science and Technological Development of the Republic of Serbia by the Grant # 179010. The authors would like to thank the university staff and students who voluntarily participated in this study.

REFERENCES

(1) Allen, M. *Misconceptions in Primary Science*, 1st ed.; Open University Press: Glasgow, UK, 2010; p 3.

(2) Stojanovska, M.; Petruševski, V. M.; Šoptrajanov, B. Study of the Use of the Three Levels of Thinking and Representation. *Nat. Math. Biotech. Sci., MASA* **2014**, *35* (1), 37−46.

(3) Novick, S.; Nussbaum, J. Junior High School Pupils' Understanding of the Particulate Nature of Matter: An Interview Study. *Sci. Educ.* **1978**, *62* (3), 273−281.

(4) Butts, B.; Smith, R. HSC Chemistry Students' Understanding of the Structure and Properties of Molecular and Ionic Compounds. *Res. Sci. Educ.* **1987**, *17* (1), 192−201.

(5) Sanger, M. J.; Greenbowe, T. J. Common Student Misconceptions in Electrochemistry: Galvanic, Electrolytic, and Concentration Cells. *J. Res. Sci. Teach.* **1997**, *34* (4), 377−398.

(6) Tongchai, A.; Sharma, M. D.; Johnston, I. D.; Arayathanitkul, K.; Soankwan, C. Developing, Evaluating and Demonstrating the Use of a Conceptual Survey in Mechanical Waves. *Int. J. Sci. Educ.* **2009**, *31* (18), 2437−2457.

(7) Pinarbasi, T.; Sozbilir, M.; Canpolat, N. Prospective Chemistry Teachers' Misconceptions about Colligative Properties: Boiling Point Elevation and Freezing Point Depression. *Chem. Educ. Res. Pract.* **2009**, *10* (4), 273−280.

DOI: 10.1021/acs.jchemed.6b00261
J. Chem. Educ. XXXX, XXX, XXX−XXX

(8) Tanahoung, C.; Chitaree, R.; Soankwan, C. Probing Thai Freshmen Science Students' Conceptions of Heat and Temperature Using Open-Ended Questions: A Case Study. *Eurasian J. Phys. Chem. Educ.* **2010**, *2* (2), 82−94.
(9) Gurel, D. K.; Eryılmaz, A.; McDermott, L. C. A Review and Comparison of Diagnostic Instruments to Identify Students' Misconceptions in Science. *Eurasia. J. Math. Sci. & Technol. Ed.* **2015**, *11* (5), 989−1008.
(10) Uzuntiryaki, E.; Geban, Ö. Effect of Conceptual Change Approach Accompanied with Concept Mapping on Understanding of Solution Concepts. *Instr. Sci.* **2005**, *33* (4), 311−339.
(11) Herrmann-Abell, C. F.; DeBoer, G. E. Using Distractor-Driven Standards-Based Multiple-Choice Assessments and Rasch Modeling to Investigate Hierarchies of Chemistry Misconceptions and Detect Structural Problems with Individual Items. *Chem. Educ. Res. Pract.* **2011**, *12* (2), 184−192.
(12) Milenković, D.; Hrin, T.; Segedinac, M.; Horvat, S. Identification of Misconceptions through Multiple Choice Tasks at Municipal Chemistry Competition Test. *J. Subj. Didact* **2016**, *1* (1), 3−12.
(13) Treagust, D. Evaluating Students' Misconceptions by Means of Diagnostic Multiple Choice Items. *Res. Sci. Educ.* **1986**, *16* (1), 199−207.
(14) Treagust, D. F. Development and Use of Diagnostic Tests to Evaluate Students' Misconceptions in Science. *Int. J. Sci. Educ.* **1988**, *10* (2), 159−169.
(15) Voska, K. W.; Heikkinen, H. W. Identification and Analysis of Student Conceptions Used to Solve Chemical Equilibrium Problems. *J. Res. Sci. Teach.* **2000**, *37* (2), 160−176.
(16) Coştu, B.; Ayas, A.; Niaz, M.; Ünal, S.; Çalık, M. Facilitating Conceptual Change in Students' Understanding of Boiling Concept. *J. Sci. Educ. Technol.* **2007**, *16* (6), 524−536.
(17) Milenković, D.; Segedinac, M.; Hrin, T. Increasing High School Students' Chemistry Performance and Reducing Cognitive Load through an Instructional Strategy Based on the Interaction of Multiple Levels of Knowledge Representation. *J. Chem. Educ.* **2014**, *91* (9), 1409−1416.
(18) Mutlu, A.; Sesen, B. A. Development of a Two-Tier Diagnostic Test to Assess Undergraduates' Understanding of Some Chemistry Concepts. *Procedia Soc. Behav. Sci.* **2015**, *174*, 629−635.
(19) Chandrasegaran, A. L.; Treagust, D. F.; Mocerino, M. 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 Reprsentation. *Chem. Educ. Res. Pract.* **2007**, *8* (3), 293−307.
(20) Gilbert, J. K. The Study of Student Misunderstandings in the Physical Sciences. *Res. Sci. Educ.* **1977**, *7* (1), 165−171.
(21) Stankov, L.; Crawford, J. D. Self-Confidence and Performance on Tests of Cognitive Abilities. *Intelligence* **1997**, *25* (2), 93−109.
(22) Pallier, G.; Wilkinson, R.; Danthiir, V.; Kleitman, S.; Knezevic, G.; Stankov, L.; Roberts, R. D. The Role of Individual Differences in the Accuracy of Confidence Judgments. *J. Gen. Psychol.* **2002**, *129* (3), 257−299.
(23) Kleitman, S.; Stankov, L. Self-confidence and Metacognitive Processes. *Learn. Individ. Differ.* **2007**, *17* (2), 161−173.
(24) Sia, D. T.; Treagust, D. F.; Chandrasegaran, A. L. High School Students' Proficiency and Confidence Levels in Displaying Their Understanding of Basic Electrolysis Concepts. *Int. J. Sci. Math. Educ.* **2012**, *10* (6), 1325−1345.
(25) Kaltakci, D.; Didis, N. Identification of Pre-Service Physics Teachers' Misconceptions on Gravity Concept: A Study with a 3-Tier Misconception Test. *Proceedings of the 6th International Conference of the Balkan Physical Union*, Istanbul, Turkey, August 22−26, 2006 10.1063/1.2733255.
(26) Eryılmaz, A. Development and Application of Three-Tier Heat and Temperature Test: Sample of Bachelor and Graduate Students. *Eurasian J. Educ. Res.* **2010**, *40*, 53−76.
(27) Peşman, H.; Eryılmaz, A. Development of a Three-Tier Test to Assess Misconceptions about Simple Electric Circuits. *J. Educ. Res.* **2010**, *103* (3), 208−222.
(28) Caleon, I. S.; Subramaniam, R. Development and Application of a Three-Tier Diagnostic Test to Assess Secondary Students' Understanding of Waves. *Int. J. Sci. Educ.* **2010**, *32* (7), 939−961.
(29) Kızılcık, H. S.; Güneş, B. Developing Three-Tier Misconception Test about Regular Circular Motion. *Hacettepe Univ. J. Educ.* **2011**, *41*, 278−292.
(30) Aykutlu, I.; Şen, A. İ. Test, Kavram Haritası Ve Analoji Kullanılarak Lise Öğrencilerinin Elektrik Akımı Konusundaki Kavram Yanılgılarının Belirlenmesi. *Educ. Sci.* **2012**, *37* (166), 275−288.
(31) Kanli, U. A Study on Identifying the Misconceptions of Pre-Service and In-Service Teachers about Basic Astronomy Concepts. *Eurasia J. Math. Sci. & Technol. Ed.* **2014**, *10* (5), 471−479.
(32) Korur, F. Exploring Seventh-Grade Students' and Pre-Service Science Teachers' Misconceptions in Astronomical Concepts. *Eurasia J. Math. Sci. & Technol. Ed.* **2015**, *11* (5), 1041−1060.
(33) Taslidere, E. Development and Use of a Three-Tier Diagnostic Test to Assess High School Students' Misconceptions about the Photoelectric Effect. *Res. Sci. Technol. Educ.* **2016**, *34* (2), 164−186.
(34) Cetin-Dindar, A.; Geban, Ö. Development of a Three-Tier Test to Assess High School Students' Understanding of Acids and Bases. *Procedia Soc. Behav. Sci.* **2011**, *15*, 600−604.
(35) Kirbulut, Z. D.; Geban, O. Using Three-Tier Test to Assess Students' Misconceptions of States of Matter. *Eurasia J. Math. Sci. & Technol. Ed.* **2014**, *10* (5), 509−521.
(36) Taagepera, M.; Noori, S. Mapping Students' Thinking Patterns in Learning Organic Chemistry by the Use of Knowledge Space Theory. *J. Chem. Educ.* **2000**, *77* (9), 1224−1229.
(37) Rushton, G. T.; Hardy, R. C.; Gwaltney, K. P.; Lewis, S. E. Alternative Conceptions of Organic Chemistry Topics Among Fourth Year Chemistry Students. *Chem. Educ. Res. Pract.* **2008**, *9* (2), 122−130.
(38) McClary, L. M.; Bretz, S. L. Development and Assessment of a Diagnostic Tool to Identify Organic Chemistry Students' Alternative Conceptions Related to Acid Strength. *Int. J. Sci. Educ.* **2012**, *34* (15), 2317−2341.
(39) Anderson, T. R.; Grayson, D. J. Improving Students' Understanding of Carbohydrate metabolism in First Year Biochemistry at Tertiary Level. *Res. Sci. Educ.* **1994**, *24* (1), 1−10.
(40) Stojanovska, M.; Šoptrajanov, B.; Petruševski, V. M. Misconceptions in the Chemistry Teaching in the Republic of Macedonia Regarding the Oxidation Reactions of Monosaccharides. *Nat. Math. Biotech. Sci.* **2013**, *34* (1−2), 27−32.
(41) Arslan, H. O.; Cigdemoglu, C.; Moseley, C. A Three-Tier Diagnostic Test to Assess Pre-Service Teachers' Misconceptions about Global Warming, Greenhouse Effect, Ozone Layer Depletion, and Acid Rain. *Int. J. Sci. Educ.* **2012**, *34* (11), 1667−1686.

DOI: 10.1021/acs.jchemed.6b00261
J. Chem. Educ. XXXX, XXX, XXX−XXX