LOOK ACADEMIC PUBLISHERS
OPEN ACCESS
INTERNATIONAL JOURNAL OF ENVIRONMENTAL & SCIENCE EDUCATION
e-ISSN: 1306-3065
2019, Vol. 14, No. 6, 345-360

Latent Profile Analysis as a Tool to Describe Students' Achievement in Entering Medicine Faculty

Helin Semilarski [1*], Anne Laius [1]

[1] *University of Tartu, ESTONIA*

* CORRESPONDENCE: helin.semilarski@ut.ee

ABSTRACT

One of the main problems in biology education is the need to know which topics in school biology are not understandable to students and which impede them to get good results in university entrance tests. The aim of the current study is to gain an overview of biological knowledge of students, who are intending to continue their education in the faculty of medicine by the end of their gymnasium studies and what are the crucial topics in which they underachieve. Based on these results, it is expected to make theoretically justified decisions what topics of biology should be emphasized in school biology and how to develop the school biology curriculum to meet the needs of competence-based biology education. The sample for this study was formed of 1017 gymnasium graduate students who fulfilled the medicine faculty entrance tests in biology (METIB) in four years (2015–2018). Based on the students' test results the tasks were categorized into four achievement levels - advanced, high, intermediate and low-performance level. The most difficult topics for students occurred to be molecules and structures and human anatomy and physiology. Research data was collected using four METIBs, which were composed of 50-item multiple-choice questions. The Latent profile analysis of test results enabled to detect three achievement profiles of students – high, moderate and low achievers. In four years of this study, the distribution of students' test scores was in correspondence with the normal distribution and was also in accordance with the Estonian National Curriculum assessment regulation system, with 15% of gymnasium graduate students achieving the highest score level and being successful in entering the faculty of medicine.

Keywords: entrance test in biology, biology test construction, core topics in biology, test evaluation, assessment results, students' achievement profiles

INTRODUCTION

Medicine is a profession which requires a very high degree of competences. It is important to identify the best possible candidates who also have a high degree of competences, including skills, knowledge, values, and attitudes. These are considered important for the success in medical field.

Medicine is one of the most popular field which has a very high demand of the students. It is one of the most recognized areas in society and therefore has very high expectations for future medical professionals. From here it is important to ensure an effective test for students and to ensure fair competition. That's why evaluation plays a major role here.

Evaluation should not only describe the state of students' achievement, but another and more meaningful goal of evaluation is to improve teaching learning process. Test, what is used to evaluate students' achievement

Article History: Received 11 June 2019 ◆ Revised 19 July 2019 ◆ Accepted 19 July 2019

needs to be valid and reliable. The method of evaluation should consist of processing the evidence needed in determining the students' level of learning and the effectiveness of the teaching. It is also important to give feedback so that topics that students do not understand too well reach to teachers. In this sense, biology and biology education can be developed through testing.

This study aims to gain an overview of the Estonian Gymnasium graduates' achievement level obtained during gymnasium studies by analysing four years' Medicine Faculty Entrance Tests in Biology (METIB) during 2015–2018 and to find out the weaknesses of students' biology knowledge that prevents the gaining of better results. Latent Profile Analysis was used to get a better overview of students' achievement levels and conclusions were drawn.

For this study, research data was collected from the 4 biology tests mentioned above, the tests composed of 50 multiple-choice (MC) questions – such testing method is used because it is cheaper and less time consuming than scoring using constructed-response questions. The results of the study have also shown the content knowledge test scores to be good predictors of success in graduating gymnasium and entering tertiary level education (Kuncel, Hezlett, & Ones, 2001).

METIB is content knowledge test that measures students' biology knowledge and understanding acquired over time in many gymnasium biology courses. The Biology curriculum identifies the main topics that students need to understand while the test identifies the conceptualization students hold about these topics and the levels of understanding through what students' progress in acquiring topics. Students needed to apply their biological knowledge and cognitive skills in order to answer the test questions. This test is used in the University of Tartu as an entrance test to the Faculty of Medicine. The last national examination in biology at gymnasium level in Estonia was in 2013 and that is one of the reasons why study authors used METIB to investigate biology education in Estonia at gymnasium level. Analysis of the entrance tests in biology is provided in this research.

The Aims of the Study and Research Questions

This study aims to investigate the difficulties that students display when answering METIB questions. Results of the investigation help educators to understand which core topics in biology are not understandable to students. If we know that, then we know what we need to improve in our biology education.

Also, one of the aims is to detect students' achievement profiles. Knowing students' profiles provides information to universities, what kind of modifications they need for their entrance test e.g. to add more high level items into the test. The importance of this study stems from importance of Biology core topics and the quality of learning of these topics as well.

In universities, students are required to take a test to be eligible for enrolment, but there have almost always been arguments about the validity of the University Entrance Test. That is one of the problems what this study aims to address.

For an analysis of this research, the following 3 research questions were posed:

RQ1. What are the achievement levels of different entrance questions?

RQ2. Which students' profiles can be distinguished based on four METIBs?

RQ3. Which core topics are difficult for students to understand, based on METIB?

The following literature review is describing all main components that are measured in this study.

LITERATURE REVIEW

Biology Education and Core Topics in Biology

The major aim of teaching is to promote the understanding of the topic being taught with a view to applying knowledge of such understanding to real life situations. Questions to METIB were developed based on core topics, which are shown in **Appendix 1**. Tests developers considered with core topics main topics in Biology Education, which should be taught in schools (Estonian National Curriculum, 2011; Brownell et al., 2014; Quinn et al., 2014). Content of the curriculum Medicine for students had been taken into consideration and it's the reason why there have been so many questions of metabolism, of molecules/organisms' structure and function and of human anatomy.

It is hypothesized that students' understanding of biology topics is related to the complexity of the biological structures and processes they involve. One of the goals of Estonian secondary school biology is that students have a good understanding of all core topics in biology. One aim of this study is to analyse students' content knowledge of those core topics in biology, to determine what content knowledge of those core topics are difficult for students to understand and which thus prevents them from applying their knowledge in different contexts. Based on these results, some theoretically justified decisions are intended to guide the enhancement of teaching and learning in school biology.

Biological Content Knowledge and Conceptual Understanding

For this conceptual framework, the biological content knowledge is focusing on 7 biology core topics (evolution; cells; metabolism; ecology; human anatomy and physiology; heredity; molecules and structures) taught according to Estonian biology curriculum in gymnasium level.

Evolution is key to understanding many other core topics for example heredity, and ecology (Tansey, Baird, Coxc, Fox, Knight, & Sears, 2013). Unfortunately, evolution is regarded as a difficult and scientifically complicated topic (Hermann, 2013; Van Dijk & Reydon, 2010; Nehm & Schonfeld, 2007). Decades of research have revealed that students regularly misunderstand what evolution is and how it occurs (Short & Hawley, 2014; Shtulman & Calabi, 2013; Sinatra et al., 2008; Sinatra et al., 2003). There has been done some research about students' understanding of evolution and results have shown that students have a poor understanding about topics of evolution (Morabito, 2010). Sager (2008) stated that due to the controversial nature of evolution, educational organizations have felt the need to explicitly state their support for the teaching of evolution. A comprehensive understanding of evolution requires competence with topics from throughout the discipline of biology as well as from other science content areas including chemistry, geology, and palaeontology (Hermann, 2013; Kampourakis & Zogza 2009).

One of the core topics in biology is cells including cell theory. Learning cell biology is an essential topic in Biology Education. Many students have misconceptions of cellular processes (Yeong, 2015) e.g. gymnasium students have problems with the topic cell division (Lukša, Radanovic, Garašic, Sertic Peric, 2016). If students have misconceptions of fundamental knowledge, then they have problems to understand biology. Lewis and her colleagues' (2010) study revealed that students' have poor understanding of the genetic relationship between cells.

Metabolism is a topic used to describe chemical reactions in the cell and in the organisms. Energy formation is included as one of the important components of metabolism.

Heredity is an important core topic in biology. There have been studies about students understanding of heredity, e.g. Kılıç & Sağlam (2014) who researched secondary school students' understanding of fundamental genetics topics.

Ecology is one core topic in biology. Students at gymnasium level have lack of understanding and misconceptions concerning fundamental topics of ecology e.g. about the greenhouse effect, energy flow (Toman, 2018).

Understanding basic human anatomy and physiology can be beneficial for all people e.g. if they need to make decisions based on their health.

There is a gap in studying secondary school students' biological understanding in the case of the main topics.

Biology Entrance Test (METIB)

In the universities, students are required to take a test to be eligible for enrolment, but there have almost always been arguments about the validity of the entrance test. That is the reason why there have been studies based on the validity of University Entrance Test (Ağazade et al., 2014; Tran et al., 2010) and study based on the history of University Entrance Test (Hatipoğlu, 2016).

The Biology test scores have been good predictors of success in graduating gymnasium and entering tertiary level education (Kuncel, Hezlett, & Ones, 2001). Evaluation should not only describe the state of students' achievement, but another goal of evaluation is to improve teaching learning process. Test, which is used to evaluate students' achievement, must be valid and reliable. The method of evaluation should consist of processing the evidence needed in determining the students' level of learning and effectiveness of the teaching. Currently, students' achievement is one of the most used measures to assess the success of school systems.

Achievement Levels and Students' Profiles

Academic achievement represents performance outcome that indicates the extent to which a student accomplished in school. Understanding educational outcomes is very important to effectively plan Biological education system. That is the reason why this study focuses on the assessment of students' biological knowledge. Academic achievement plays a major role in life. Higher achievement in science is important for students to prepare them for university, and life in our changing world (Mullis *et al*., 2012). Mainly, there are used 3 achievement levels in science: high, average, and low (Abed, 2016). Sungur and her colleagues (2006) measured students' academic achievement by multiple choice questions.

Kablan & Kaya (2013) study focused on correlations between different items and students' science achievement. In this study, they used 3 different items: knowing, applying and reasoning. Item's difficulty can be subjective, which should be assessed by the item doers (Li & Belkin, 2008).

Test takers could be presented in different profile groups in which each profile has its specific properties. In this study, we seek to find a way to categorise students into different profile groups based on their biology test scores, while Brunner and his colleagues (2013) presented in their study models based on mathematics achievement.

METHODOLOGY

Data Collection Procedures

The data for this study were collected from four years (2015–2018) METIB tests. The students had 120 minutes to answer 50 MC questions.

Participants

The participants for this study are gymnasium graduate students, varying from age 18 to 25. In total, 278 (70 males and 208 females) students were involved in 2015; 270 (50 males, and 220 females) in 2016, and 251 (62 males, and 188 females) in 2017; 218 (56 males, and 162 females) in 2018 entrance tests, thus the total number of participants was 1017.

Instrument

Data was collected from METIBs (2015–2018), these were multiple-choice tests, each comprising of 50 questions. Closed-response items (such as multiple-choice questions) are popular for their efficiency and reliability. The tests in biology have 50 questions for 1 point (50 points total). In this study, there were used 4 tests (METIB 2015–2018) and they were authors' own constructions.

Multiple-choice (MC) test has become popular at every level of education. Multiple-choice tests have many advantages e.g. they are easy to score, allow more content to be covered and offer better objectivity in grading (Butler, 2018). Although tests are needed for teachers to assess students understanding, they are good for students also for learning (Pan & Rickard, 2018). Answering to different test questions can help to increase students understanding. One misconception about multiple-choice items is that they measure only factual knowledge. Aiken (1982) introduced in his study five types of items that can be used to assess higher-order educational aims also.

Created instrument followed following criterions:

- Instrument was conducted based on Estonian Biology Curriculum (questions covered the main topics) and the items were meant to measure the students' different achievement levels (Estonian curriculum, 2011).
- 50 MC items measure biological conceptual understanding and cognitive competences.

The aim of the test is to differentiate students who have good competences from the students who don't have so good competences.

Table 1. METIBs reliabilities and validities 2015–2018

Year	Reliability (Cronbach α)	Validity	Validation method used
2015	0.80	Content validity	Expert opinion method: two biologists, three biology teachers, two educational researchers. In total seven independent experts in the field of science education.
2016	0.86	Construct validity	Analysis of Estonian gymnasium biology curriculum to ensure tha items are valid in terms of expected learning outcomes.
2017	0.84		
2018	0.87		

Experts evaluated whether questions in the test are following Estonian biology syllabus. Regarding METIB, the assessment of the competences is closely related to the subjects taught at school. Also, item difficulties were measured during every year. Items' difficulties varied from 0.14–0.94 (**Appendix 2**).

Validity and Reliability

Reliability of tests was found in each year (**Table 1**). The values of Cronbach's alpha (α = 0.80 in 2015, α=0.87 in 2016, α=0.84 in 2017, and α=0.87 in 2018) have indicated a high internal consistency of tests. The reliability of the instruments was determined to be acceptable using Cronbach alpha (0.80–0.87). The validity of the instruments (4 years METIBs) and the methodology used was determines as shown in **Table 1**. Items in the tests varied over four years. That is the reason why instruments validated separately every year. All seven independent experts were the same during all four validations. 3 experts categorized and grouped questions based on the topics.

Data Analysis

Quantitative data were collected for this study and these were analysed with IBM SPSS Statistics 25 and Mplus. Differences were calculated using the Wilcoxon test in SPSS to compare differences between the mean scores for the four years' tests. Latent profile analysis was performed with the Mplus Editor program (Muthén & Muthén, 2012). Latent profile analysis is a person-oriented mixture modelling analysis (Williams & Kibowski, 2016) that was used to distinguish METIB students' profiles based on the achievement levels of the test. For the evaluation of the overall model fit, three different fit indices were used.

A qualitative content analysis approach was used to analyse and categorize the test questions. Content analysis has been defined as a systematic technique for compressing many words of text into fewer content categories based on the coding rules (Krippendorff, 1980; Weber, 1990). For this study, the qualitative content analysis (Schreier, 2012) began with the creation of a coding frame, followed by analysing questions of the test. Content analysing was used to categorize questions, based on topics. The categorized questions were validated by three experts (two educational researchers and one biology teacher).

RESULTS

Achievement Levels based on the Item Difficulties

Analysis of METIB' results showed that students mean results answering to the questions vary highly (respectively 24% to 88%). Mean results were expressed in percentages of the maximum possible outcome and based on that three achievement levels were created (easy items, medium easy items, medium difficult items, and highly difficult items). Additionally, achievements levels (according to item difficulty, which was decided according to the mean score) were defined:

1) advanced performed items (80–100%);
2) high performed items (60–79%):
3) intermediate performed items (40–59%);
4) low performed items (0–39%).

Each item was evaluated occurring to the achievement level of students' results. According to the mean results of the items four achievement level groups differed (advanced, high, intermediate, and low performed items) and bonded significantly ($p<0.05$) with χ^2 test (**Table 2**). Based on the chi-square test ($p<0.05$) it is

Table 2. Four item achievement levels of 4 METIBs 2015–2018

Achievement level	**Year**	**Number of items**	**Mean results of items** (SD)	**Level of achievement** (%)	**X^2 test p**
Advanced performed items (scores are equal to or higher than 80%)	2015	9	**0.86** (0.34)	86	0.0003*
	2016	8	**0.83** (0.37)	83	
	2017	9	**0.85** (0.36)	85	
	2018	1	**0.83 (0.45)**	83	
High performed items (60% to less than 79%)	2015	17	**0.70** (0.46)	70	0.0000*
	2016	19	**0.70** (0.45)	70	
	2017	19	**0.69** (0.46)	69	
	2018	15	**0.68** (0.46)	68	
Intermediate performed items (40% to less than 59%)	2015	15	**0.52** (0.50)	52	0.0001*
	2016	14	**0.48** (0.50)	50	
	2017	14	**0.49** (0.50)	49	
	2018	24	**0.48** (0.50)	48	
Low performed items (0% to less than 39%)	2015	9	**0.30** (0.45)	30	0.0000*
	2016	9	**0.30** (0.46)	30	
	2017	8	**0.31** (0.46)	31	
	2018	10	**0.32** (0.46)	32	

* Statistical significance at the level $p<0.05$

concluded that there is a significant relationship between achievement levels and the mean results of the items.

Based on the 2015th METIB, there were 9 easy items, 17 medium easy items, 15 medium difficult items and 9 highly difficult items. Based on the 2016th METIB, there were 8 easy items, 10 medium easy items, 22 medium difficult items, and 10 highly difficult items. Based on the 2017th METIB, there were 9 easy items, 19 medium easy items, 14 medium difficult items, and 9 highly difficult items. Based on the 2018th METIB, there were 1 easy item, 15 medium easy items, 24 medium difficult items, and 10 highly difficult items.

Numbers of items at every achievement level have been basically the same over the 4 years. Every year mean results of items have been similar within all achievement levels e.g. the mean result differed between 0.69 and 0.70 in medium easy items.

Students' Achievement Profiles based on 2015—2018 METIB

The highest score was 50 in each four years, but nobody reached that score. As a result of the study three students' achievement profiles were detected according to their test results by Semilarski and Laius (2019): (1) High Achievers (students' exhibit advanced performance when tested on skills and concepts); (2) Moderate Achievers (students' exhibit some appropriate performance); (3) Below Moderate Achievers (students don't exhibit appropriate performance); and (4) Low Achievers (students who don't exhibit any appropriate performance).

One of the research questions was "Which students' profiles can distinguish based on four METIBs". Latent profile analysis was used to distinguish student groups based on their response profiles. The variables included in the analysis were based on items which have different difficulties (highly easy items, lowly easy items, lowly difficult items, highly difficult items). Different years were also included, so student profiles were compared based on 2015, 2016, 2017 and 2018.

The best model of 2015th METIB students was chosen based on a combination of model results and fit statistics (Ram & Grimm, 2009). The 3-profile model was theoretically sensible and showed the best fit based on AIC, BIC, entropy (shows classification quality), and A-LMRT (fit statistics of 2- and, 3-profile model are shown in **Table 3**). It was impossible to make students 4- and 5-profile models with this data set.

Table 3. Fit statistics (Loglikelihood, AIC, BIC A-LMRT, entropy) of 2- and 3-profile models of METIB 2015

Model fit Statistics	2-profile model	3-profile model	4-profile model	5-profile model
Loglikelihood (H0) value	-7920.803	-7800.465	X	X
Akaike (AIC)	16143.607	16004.930	X	X
Bayesian (BIC)	16691.378	16737.710	X	X
Entropy	0.912	0.865	X	X

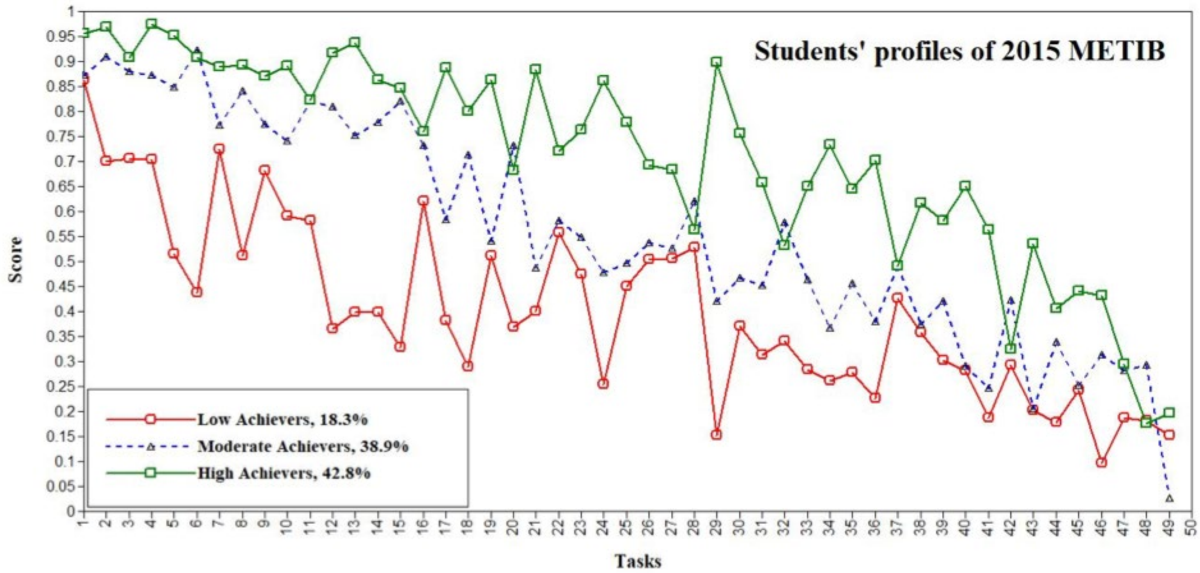

Figure 1. The 2015th METIB students' 3-profile model

Table 4. Fit statistics (Loglikelihood, AIC, BIC A-LMRT, entropy) of 2-, 3-, and 4-profile models of METIB 2016

Model fit Statistics	2-profile model	3-profile model	4-profile model
Loglikelihood (H0) value	-7582.107	X	-6889.899
Akaike (AIC)	15466.214	X	14285.799
Bayesian (BIC)	16009.575	X	15196.199
Entropy	0.984	X	0.995

Vuong-Lo-Mendell-Rubin likelihood ratio test for 2 classes showed that p-value was 0.000, and it shows that data doesn't fit with the model. Same test showed that for 3 classes p-value was 0.609, showing that data fits with model (Fisher, 1925). The 2015th METIB students' 3-profile model is presented in **Figure 1**. Based on the results there were 50 low, 109 moderate, and 119 high achievers.

High Achievers had highest results mainly in all 50 items. Low achievers had lowest results mainly in all items. Only in the last item, they performed better than moderate achievers. There were 51 low achievers, 108 moderate achievers, and 119 high achievers in the 2015th METIB.

The best model of 2016th METIB students was chosen based on a combination of model results and fit statistics (Ram & Grimm, 2009). The 2-profile model was theoretically sensible and showed the best fit based on Akaike's information criterion (AIC), Schwarz's information criterion (BIC), entropy (shows classification quality), and A-LMRT (fit statistics of 2, 3, 4-profile model are shown in **Table 4**). It was impossible to make students' 3-, 4- and 5-profiles with this data set.

Vuong-Lo-Mendell-Rubin likelihood ratio test for 2 classes showed that data fits with the model. The 2016th METIB students' 3-profile model is presented in **Figure 2**. Based on the results there were 63 low, and 207 high achievers. This year the moderate profile of students was missing.

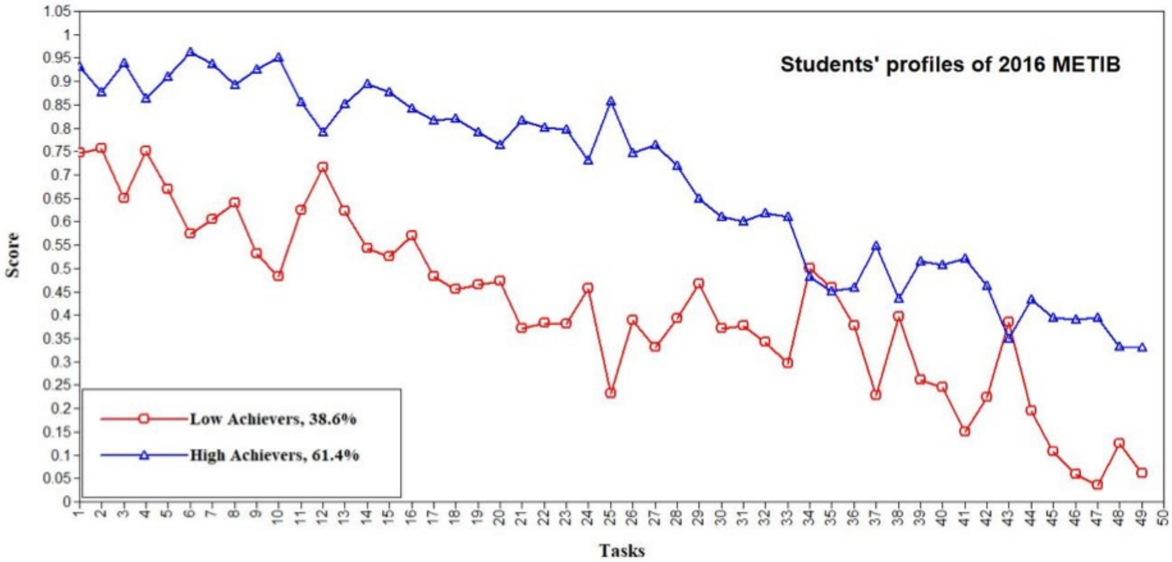

Figure 2. The 2016th METIB students' 2-profile model

Table 5. Fit statistics (Loglikelihood, AIC, BIC A-LMRT, entropy) of 2-, 3-, 4-, and 5-profile models of METIB 2017

Model fit Statistics	2-profile model	3-profile model	4-profile model	5-profile model
Loglikelihood (H0) value	-7085.276	**-6880.637**	-6810.854	X
Akaike (AIC)	14472.552	**14165.273**	14127.709	X
Bayesian (BIC)	15004.293	**14876.608**	15018.638	X
Entropy	0.926	**0.931**	0.951	X

The best model of 2017th METIB students was chosen based on a combination of model results and fit statistics (Ram & Grimm, 2009). The 3-profile model was theoretically sensible and showed the best fit based on AIC, BIC, entropy (shows classification quality), and A-LMRT (fit statistics of 2, 3, and 5-profile model are shown in **Table 5**). It was impossible to make students' 5-profile model, with this data set.

Latent Profile Analysis couldn't make 5-profile model of METIB 2017. Vuong-Lo-Mendell-Rubin likelihood ratio test for 2 classes showed that p-value was 0.000, and it shows that data doesn't fit with the model. The same test showed that for 3 classes p-value was 0.0811 and for 4 classes 0.712. 3-classes model fits better with data. The 2017th METIB students 3- profiles are presented in **Figure 3**. Based on the results there were 68 low, 112 moderate, and 70 high achievers.

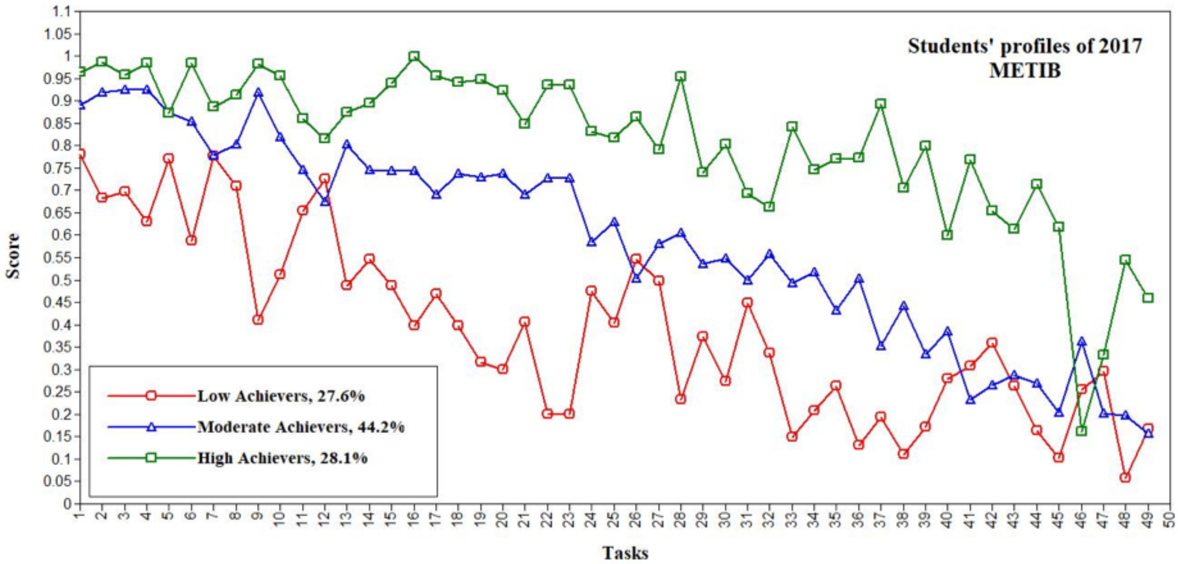

Figure 3. The 2017th METIB students' 3-profile models

Table 6. Fit statistics (Loglikelihood, AIC, BIC A-LMRT, entropy) of 2-, 3-, 4-, and 5-profile models of METIB 2018

Model fit Statistics	2-profile model	3-profile model	4-profile model	5-profile model
Loglikelihood (H0) value	-7370.558	**-6667.653**	-6810.854	-6566.496
Akaike (AIC)	13870.136	**13739.307**	14127.709	13740.993
Bayesian (BIC)	14381.195	**14422.975**	15018.638	14769.879
Entropy	0.944	**0.947**	0.951	0.948

High Achievers had highest results mainly in all 50 items. Low Achievers had lowest results mainly in all items. Only in one item high achievers' performance was the lowest.

The best model of 2018th METIB students was chosen based on a combination of model results and fit statistics (Ram & Grimm, 2009). The 3-profile model was theoretically sensible and showed the best fit based on AIC, BIC, entropy (shows classification quality), and A-LMRT (fit statistics of 2, 3, and 5-profile model are shown in **Table 6**).

The 2018th METIB students 3- profiles are presented in **Figure 4**. Based on the results there were 68 low, 112 moderate, and 70 high achievers.

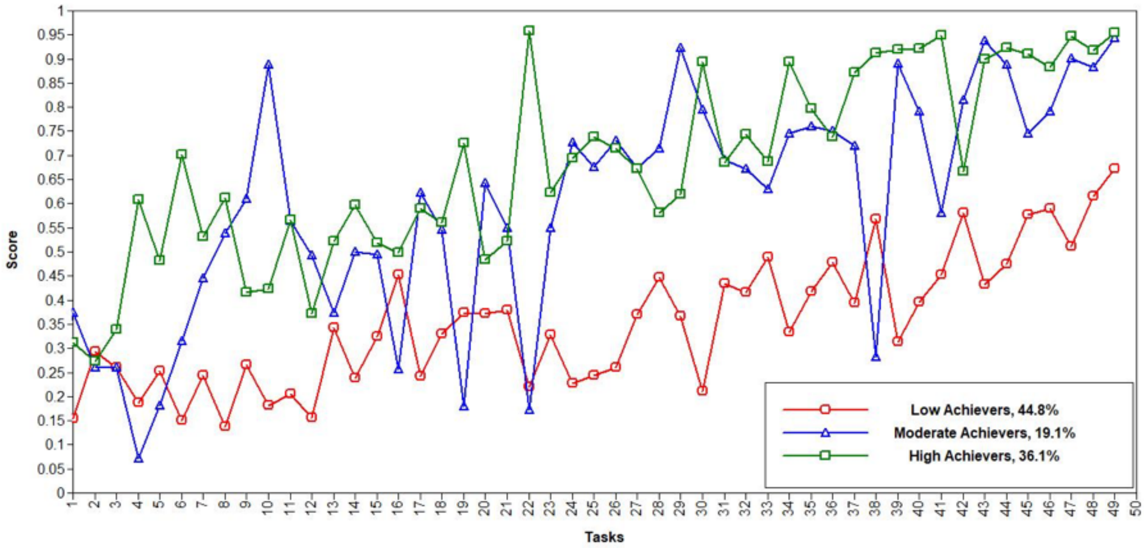

Figure 4. The 2018th METIB students' 3-profile models

Table 7. METIB questions categorized based on included core topic and percentage of correct answers

Core topic	No. of items in METIB	Percentage of correct answers (%)
Ecology	15	64
Cells	38	63
Evolution	10	59
Heredity	35	58
Metabolism	28	52
Molecules and structures	32	50
Human Anatomy and Physiology	42	47

Students' Understanding of Different Core Topics

Third research question focused on different core topics in biology. Content analysis was used to categorise items based on with which core topic it is connected. To do that all four year METIB' items were analysed together, the total number of items is 200. Seven experts determined that tests items included 7 core topics: heredity, ecology, evolution, molecules and structures, metabolism, cells and human anatomy and physiology (**Table 7**).

Human anatomy and physiology was assessed by the biggest number of items as this topic is very important for medicine students, but the results showed that this was the most difficult topic for students.

CONCLUSION AND DISCUSSION

The major aim of teaching is to promote the understanding of the topic being taught with a view to applying knowledge of such understanding to real life situations. This study focused to evaluate students' biological knowledge through core topics in biology. The aim of this study was to gain an overview of the Estonian Gymnasium graduates' achievement level obtained during gymnasium studies by analysing four METIBs (during 2015–2018).

The first research question addressed in this paper focused on determining the achievement levels of different entrance test questions. Results showed that students mean results answering the questions vary highly (respectively 24% to 88%). There are many studies about how to categorise different items (Li & Belkin, 2008), in our study we categorised items based on their difficulty. According to item difficulty (which was decided according to the mean score of the items), there were 4 different categories of items: advanced performed items, high performed items, intermediate performed items, and low performed items.

Student achievement is most relevant indicator of teaching and learning effectiveness. Based on student achievement we can determine, if student has learned effectively or not. The second research question in this paper focused on determining the students' profiles based on four METIBs. Latent Profile Analysis was used to distinguish different students' profiles. Results showed that we can distinguish students into profiles, based on their' achievement. Based on 4 METIBs we can distinguish 3 students' profiles: high, moderate and low achievers. Students' biology achievement levels don't vary significantly across the four years. Results showed that there were more low achievers than high achievers.

One of the aims was to find out the weaknesses of students' biology knowledge that prevents the gaining of better results. One of the goal of Estonian secondary school biology curriculum is that students have good understanding about all core topics of biology. The final research question addressed in this paper focused on determining, which core topics in biology are difficult for students to understand, based on METIB. Test results showed that there were different core topics that are difficult to understand, but the level of achievement depended on the complexity of the items. This study and Kablan & Kaya (2013) found similar results that achievement decreased as questions became more complex.

The results of the examination met the requirements of the assessment system and were suitable to assess the candidates' biological knowledge. In all four years, the distribution of students' entrance examination scores followed a normal distribution and in accordance to the assessment regulation system, 15% of students achieved the highest score level and were successful in the entrance to the faculty of medicine.

LIMITATION OF THE STUDY

The cohort of students who participated in this study were drawn from only 4 years and because of this there may be bias in the conclusions drawn from this study. The instrument was consisted only of multiple-choice questions, which didn't allow measuring all aspects of biological literacy. Despite to these limitations, the findings of this study are still thought to represent a valuable contribution to biology education. Multiple-choice items allow valid and reliable measurement of wide range of curriculum outcomes, but are less satisfactory for assessing some outcomes such as problem solving, decision making and reasoning and creativity.

IMPLICATIONS OF THE TEACHING OF BIOLOGY

The results of this study could be used by biology educators in further preparing the students for entering the tertiary level of education in the field of medicine.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Helin Semilarski – University of Tartu, Estonia.

Anne Laius – University of Tartu, Estonia.

REFERENCES

Abed, H. O. (2016). Drama-Based Science Teaching and Its Effect on Students' Understanding of Scientific Concepts and Their Attitudes towards Science Learning. *International Education Studies*, *9*(10). https://doi.org/10.5539/ies.v9n10p163

Ağazade, S. A., Caner, H., Hasipoğlu, N. H., & Civelek, H. A. (2014). Turkish University Entrance Test and academic achievement in undergraduate programs: a criterion related validity study. *Procedia – Social and Behavioral Sciences*, *116*, 4582–4590. https://doi.org/10.1016/j.sbspro.2014.01.990

Aiken, R. A. (1982). Writing Multiple-Choice Tasks to Measure Higher-Order Educational Objectives. *Educational and Psychological Measurement*, *42*, 803–806. https://doi.org/10.1177/001316448204200312

Brownell, S., Freeman, S., Wenderoth, P. M. and Crowe, J. A. (2014). BioCore Guide: A Tool for Interpreting the Core Concepts of Vision and Change for Biology Majors. *Life Sciences Education*, *13*, 200–211. https://doi.org/10.1187/cbe.13-12-0233

Brunner, M., Gogol, M., K., Sonnleitner, P., Keller, U., Krauss, S. & Preckel, F. (2013). Gender differences in the mean level, variability, and profile shape of student achievement: Results from 41 countries. *Intelligence*, *41*, 378–395. https://doi.org/10.1016/j.intell.2013.05.009

Butler, C. A. (2018). Multiple-Choice Testing in Education: Are the Best Practices for Assessment Also Good for Learning. *Journal of Applied Research in Memory and Cognition*, *7*, 323–331. https://doi.org/10.1016/j.jarmac.2018.07.002

Fisher, R. A. (1925). Theory of Statistical Estimation. *Proceeding of the Cambridge Philosophical Society*, *22*, 700–725. https://doi.org/10.1017/S0305004100009580

Hasipoğlu, Ç. (2016). The Impact of the University Entrance Exam on EFL Education in Turkey: Pre service English Language Teachers' Perspective. *Procedia – Social and Behavioral Sciences*, *232*, 136–144. https://doi.org/10.1016/j.sbspro.2016.10.038

Hermann, R. S. (2013). High school biology teachers' views on teaching evolution: implications for science teacher educators. *Journal of Science Teacher Education*, *24*, 597–616. https://doi.org/10.1007/s10972-012-9328-6

Kablan, Z., & Kaya, S. (2013). Science achievement in TIMSS cognitive domains based on learning styles. *Egitim Arastirmalari – Eurasian Journal of Educational Research*, *53*, 97–114. https://doi.org/10.14689/ejer.2013.53.6

Kılıç, D., & Sağlam, N. (2014). Students' understanding of genetics concepts: the effect of reasoning ability and learning approaches. *Journal of Biological Education*, *48*(2), 63–70. https://doi.org/10.1080/00219266.2013.837402

Krippendorff, K. (1980). *Content Analysis: An Introduction to Its Methodology*. Newbury Park, CA: Sage.

Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2001). A comprehensive meta-analysis of the predictive validity of the Graduate Record Examinations: *Implications for graduate student selection and performance. Psychological Bulletin*, *127*, 162–181. https://doi.org/10.1037/0033-2909.127.1.162

Lewis, J., Leach J. & Wood-Robinson, C. (2000). What's in a cell? — young people's understanding of the genetic relationship between cells, within an individual. *Journal of Biological Education*, *34*(3), 129–132. https://doi.org/10.1080/00219266.2000.9655702

Li, Y., & Belkin, N.J. (2008). A faceted approach to conceptualizing items in information seeking. *Information Processing & Management*, *44*, 1822–1837. https://doi.org/10.1016/j.ipm.2008.07.005

Lukša, Ž., Radanovic, I., Grašic, D., & Sertic Peric, M. (2016). Misconceptions of Primary and High School Students Related to the Biological Concept of Human Reproduction, Cell Life Cycle and Molecular Basis of Heredity. *Journal of Turkish Science Education, 13*(3), 143–160.

Morabito, P. N., Catley, M. K., & Novick, R. L. (2010). Reasoning about evolutionary history: Postsecondary students' knowledge of most recent common ancestry and homoplasy. *Journal of Biological Education*, *44*(4), 166–174. https://doi.org/10.1080/00219266.2010.9656217

Mullis, I., V. S., Martin, M., O., Foy, P., & Arora, A. (2012). *TIMSS 2011 International Results in Mathematics.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Muthén, L. K., & Muthén, B. O. (1998–2012). Mplus User's Guide. Seventh Edition. Los Angeles, CA: Muthén & Muthén.

National curricula for upper secondary schools. (2011). Appendix 4: Natural Science. Retrieved from https://www.riigiteataja.ee/aktilisa/1140/1201/1001/VV1_lisa4.pdf

National Research Council. (2012). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas. Committee on a Conceptual Framework for New K-12 Science Education Standards. Board on Science Education, Division of Behavioral and Social Sciences and Education*. Washington, DC: National Academies Press.

Nehm, H., R., & Schonfeld, S., I. (2007). Does Increasing Biology Teacher Knowledge of Evolution and the Nature of Science Lead to Greater Preference for the Teaching of Evolution in Schools? *Science Teacher Education*, *18*, 699–723. https://doi.org/10.1007/s10972-007-9062-7

Pan, S. C., & Rickard, T. C. (2018). Transfer of test-enhanced learning: Meta-analytic review and synthesis. *Psychological Bulletin*, *144*, 710–756. https://doi.org/10.1037/bul0000151

Pesina, S., & Solonchak, T. (2015). Concept in Cognitive Linguistics and Biocognitive Science. *Social and Behavioral Sciences*, *192*, 587–592. https://doi.org/10.1016/j.sbspro.2015.06.100

Quinn, H., Schweingruber, H., & Keller, T. (Eds.). (2011). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Committee on Conceptual Framework for the New K-12 Science Education Standards; National Research Council, USA.

Sager, C. (2008). *Voices of evolution*. Berkley, CA: The National Center for Science Education, Inc.

Schreier, M. 2012. *Qualitative Content Analysis in Practice*. London: Sage.

Semilarski, H., & Laius, A. (2019). A complex instrument for measuring the components of gymnasium students' biological literacy. *EDULEARN19 Proceedings: 11th International Conference on Education and New Learning Technologies*, 1st-3rd July, 2019. Palma, Mallorca, SPAIN: iated.org/edulearn, 6285−6293.

Short, S. D., & Hawley, P. H. (2014). The effects of evolutionary education: Examining attitudes toward and knowledge of evolution in college courses. *Evolutionary Physiology*, *13*(1), 67–88. https://doi.org/10.1177/147470491501300105

Shtulman, A., & Calabi, P. (2013). Tuition vs. intuition: Effects of instruction on naïve theories of evolution. *Merrill-Palmer Quarterly Journal of Developmental Physiology*, *59*, 141–167. https://doi.org/10.1353/mpq.2013.0010

Sinatra, G. M., Brem, S. K., & Evans, E. M. (2008). Changing minds? Implications of conceptual change for teaching and learning about biological evolution. *Evolution: Education Outreach*, *1*, 189–195. https://doi.org/10.1007/s12052-008-0037-8

Sinatra, G. M., Southerland, S. A., McConaughy, F., & Demastes, J. (2003). Intentions and beliefs in students' understanding and acceptance of biological evolution. *Journal of Research in Science Teaching*, *40*(5), 510–528. https://doi.org/10.1002/tea.10087

Sungur, S., Takkaya, C., & Geban, Ö. (2006). Improving achievement through problem-based learning. *Journal of Biological Education*, *40*(4), 155–160. https://doi.org/10.1080/00219266.2006.9656037

Surman, J., Stráner, K., & Haslinger, P. (2014). Nomadic concepts in the history of biology. *Studies in History and Philosophy of Biological and Biomedical Sciences*, *48*, 127–129. https://doi.org/10.1016/j.shpsc.2014.08.001

Tansey, J. T., Baird, T., Cox, M. M., Fox, K. M., Knight, J., & Sears, D. (2013). Foundational Concepts and Underlying Theories for majors "Biochemistry and Molecular Biology". *Biochemistry and Molecular Biology Education*, *41*(5), 289–296. https://doi.org/10.1002/bmb.20727

Toman, U. (2018). An Investigation into the Learning of Ecological Concepts. *European Journal of Educational Research*, *7*(3), 631–638. https://doi.org/10.12973/eu-jer.7.3.631

Tran, P. H., Griffin, P., & Nguyen, C. (2010). Validating the university entrance English test to the Vietnam National University: A conceptual framework and methodology. *Procedia – Social and Behavioral Sciences, 2*, 1295–1304. https://doi.org/10.1016/j.sbspro.2010.03.190

Van Dijk, E. M., & Reydon, T. A. (2010). A conceptual analysis of evolutionary theory for teacher education. *Science Education*, *19*(6–8), 655–677. https://doi.org/10.1007/s11191-009-9190-x

Weber, R. P. (1990). *Basic Content Analysis*. Newbury Park, CA. https://doi.org/10.4135/9781412983488

Williams, G. A., & Kibowski, F. (2016). Latent Class Analysis and Latent Profile Analysis. In: L. A. Jason, D. S. Glenwick (Ed.). *Handbook of methodological approaches to community-based research: Qualitative, quantitative, and mixed methods*. Oxford University Press. https://doi.org/10.1093/med:psych/9780190243654.003.0015

Yeong, M., F. (2015). Using Primary Literature in an Undergraduate Assignment: Demonstrating connections among cellular processes. *Journal of Biological Education*, *49*(1), 73–90. https://doi.org/10.1080/00219266.2014.882384

Zohar, A., & Nemet, F. (2001). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching*, *39*, 35–62. https://doi.org/10.1002/tea.10008

APPENDIX 1

Comparison between different main topics/core concepts in biology at secondary school level

Biology topics included in the Biology Curriculum in Estonia (Estonian National Curriculum, 2011)		Core and component ideas in the biology (Quinn *et al.*, 2014)		5 core concepts for undergraduate biology education (Brownell *et al.*, 2014)	
Topic	**Sub-topics**	**Core idea**	**Sub-component ideas**	**Core concept**	**Overarching principles:**
Cells	Research fields in Biology Chemical Composition of Organisms Cell (Cell Theory) Cell Diversity	From Molecules to Organisms: Structures and Processes	Structure and Function Growth and Development of Organisms Organization for Matter and Energy Flow in Organisms Information Processing	Structure and function	E.g. The structure of cell and molecule
Organisms	Organisms' metabolism Development of Organisms Function and regulation in the Human body	Ecosystems: Interactions, Energy, and Dynamics	Interdependent Relationships in Ecosystems Cycles of Matter and Energy Transfer in ecosystems Ecosystem dynamics, Functioning and Resilience Social Interactions and Group Behaviour	Transformations of energy and matter	E.g. Natural selection leads to the evolution of efficient use of resources within constraints
Molecular biology	Molecular Mechanisms of Biological Processes Viruses and Bacteria Heredity and Variability	Heredity: Inheritance and Variation of Traits	Inheritance of Traits Variation of Traits	Systems	E.g. Organisms have complex systems that allow them to respond to changes in environment
Evolution and ecology	Biological Evolution Ecology Environmental Biology	Biological evolution: Unity and Diversity	Evidence of Common Ancestry and diversity Natural Selection Adaptation Biodiversity and Humans	Evolution	E.g. Species evolve over time Natural selection Gene flow Genetic drift
				Information flow	E.g. Cells/organs have multiple mechanisms to perceive and respond to changing environmental conditions.

APPENDIX 2

Description of items and item difficulty of METIB

Item No.	Exam year	Topic	Content of items	Item difficulty
13	2015	Molecules and structures	Nucleotides	0.80
12	2017	Molecules and structures	DNA molecule	0.70
12	2015			0.67
2	2018	Molecules and structures	Peptide molecule structure	0.30
2	2017			0.26
3	2018	Molecules and structures	Cytokinesis	0.66
4	2017			0.70
45	2018	Molecules and structures	Transcription	0.62
23	2015	Molecules and structures	Relationship between enzyme and DNA	0.50
30	2018	Molecules and structures	Replication of semi-conservative process	0.49
46	2017			0.42

Item No.	Exam year	Topic	Content of items	Item difficulty
8	2015	Molecules and structures	Biological taxa	0.48
6	2018	Molecules and structures	Meiotic division in one diploid animal cell	0.46
7	2017			0.54
9	2015			0.59
8	2018	Molecules and structures	Molecules in intercellular substance	0.40
9	2017			0.37
7	2015			0.37
31	2018	Molecules and structures	DNA parts labelling on a drawing	0.39
47	2017			0.42
21	2015	Molecules and structures	Importance of Sulphur in organisms	0.33
5	2018	Molecules and structures	Translation processes on a diagram	0.28
6	2017			0.26
32	2015	Metabolism	Identification of amylase	0.83
48	2017	Metabolism	Glycose producing	0.72
34	2018	Metabolism	Calvin cycle in the chloroplast	0.71
11	2018	Metabolism	Photosynthetic light reactions in the Calvin cycle	0.69
15	2017			0.66
14	2015			0.49
21	2017	Metabolism	Krebs' cycle	0.62
15	2015			0.51
16	2015	Metabolism	Light and dark reactions of photosynthesis in a chloroplast	0.62
15	2018	Metabolism	Useless light	0.58
20	2017			0.48
32	2018	Metabolism	Photosystem II	0.57
49	2017			0.54
35	2015	Metabolism	Glucagon	0.51
33	2018	Metabolism	Enzyme lactase	0.50
11	2017	Metabolism	Glycolysis	0.50
11	2015			0.58
38	2018	Metabolism	activity of the enzyme	0.43
12	2018	Metabolism	anaerobic glycolysis	0.38
16	2017			0.55
22	2015	Metabolism	Chlorophyll absorption	0.32
44	2018	Metabolism	Plants eaters	0.19
1	2018	Cells	Characteristics of living things	0.82
1	2017			0.78
1	2015			0.88
40	2017	Cells	Oogenesis and spermatogenesis	0.80
41	2015			0.90
4	2018	Cells	Multicellular organisms	0.78
5	2017			0.73
10	2018	Cells	Plant and animal cell	0.74
13	2017			0.68
10	2015			0.77
3	2017	Cells	Cell Theory	0.73
2	2015			0.72
7	2018	Cells	Organelles order in exocytosis	0.72
8	2017			0.67
5	2015			0.77
4	2015	Cells	Interphase processes	0.64
43	2018	Cells	Interphase	0.62
13	2018	Cells	Mitochondria	0.59
18	2017			0.62
19	2015			0.64
3	2015	Cells	Phase of mitosis	0.60
22	2018	Cells	Fertilisation of the plants (egg and pollen)	0.56
35	2017			0.29
47	2018	Cells	Cell Theory (cells existence)	0.53
20	2018	Cells	Unicellular algae	0.45
31	2017			0.52
9	2018	Cells	Metaphase	0.40
10	2017			0.40
6	2015			0.50
50	2017	Human anatomy and physiology	Down syndrome	0.82
17	2017	Human anatomy and physiology	Aerobic and anaerobic respiration in human body	0.73
17	2015			0.74
46	2015	Human anatomy and physiology	Embryogenesis	0.63
33	2017	Human anatomy and physiology	Insulin production	0.62
37	2018	Human anatomy and physiology	Half-moon valves	0.61
45	2015	Human anatomy and physiology	Placenta function	0.58
27	2018	Human anatomy and physiology	Blood clotting	0.57
42	2017			0.62
44	2015			0.60

Item No.	Exam year	Topic	Content of items	Item difficulty
46	2018	Human anatomy and physiology	Colon function	0.54
50	2015	Human anatomy and physiology	Blood vessels	0.47
25	2018	Human anatomy and physiology	Human digestive system	0.46
39	2017			0.42
23	2018	Human anatomy and physiology	Positive feedback	0.44
36	2017			0.87
38	2015			0.57
35	2018	Human anatomy and physiology	Passive immunity	0.43
29	2018	Human anatomy and physiology	Cholesterol in human body	0.42
45	2017			0.48
26	2018	Human anatomy and physiology	Blood flow through lung vein	0.39
41	2017			0.40
43	2015			0.27
48	2018	Human anatomy and physiology	Muscles and inhalation	0.39
47	2015	Human anatomy and physiology	Dialysis apparatus	0.36
24	2015	Human anatomy and physiology	Human Genome project	0.35
42	2015	Human anatomy and physiology	Bacterial infection	0.34
14	2018	Human anatomy and physiology	Phosphate in human body	0.31
19	2017			0.36
44	2017	Human anatomy and physiology?	Hypothermia	0.28
49	2015			0.22
28	2018	Human anatomy and physiology?	Mesoderm	0.25
43	2017			0.24
24	2017	Heredity	Skin pigmentation	0.89
24	2015			0.12
14	2017	Heredity	Amino acids order based on the mRNA	0.87
26	2017	Heredity	Inheritance patterns of human blood groups (AB0 types)	0.86
29	2015			0.94
23	2017	Heredity	Estonian gene pool	0.84
25	2017	Heredity	Human cloning	0.81
28	2015			0.86
37	2015	Heredity	Genetic variety in taxa	0.80
26	2015	Heredity	Genetics of the Rhesus blood group system	0.74
17	2018	Heredity	Heredity of Haemophilia	0.73
27	2017			0.76
30	2015			0.78
48	2015	Heredity	Flower reproduction	0.68
18	2015	Heredity	The universality of the genetic code	0.67
16	2018	Heredity	Allele and gene	0.66
22	2017			0.73
27	2015	Heredity	Information from genealogy tree	0.64
41	2018	Heredity	Inheritance patterns in solving cross problem set	0.61
39	2018	Heredity	Human blood clotting factor	0.48
36	2018	Heredity	Human skin colour and heredity	0.41
50	2018	Heredity	Genotype	0.32
40	2018	Heredity	Inherited traits	0.29
32	2017	Ecology	Global warming	0.88
29	2017	Ecology	Consequences, when the Earth's average temperature continues to rise	0.81
33	2015			0.81
28	2017	Ecology	Energy transforming	0.64
31	2015			0.83
49	2018	Ecology	Mathematical assignment about tropical levels and biomass production	0.63
36	2015	Ecology	Food network	0.51
18	2018	Ecology	The growth of the biomass of consumers	0.46
19	2018	Ecology	Trophic levels of the ecosystem	0.48
30	2017			0.46
40	2015	Ecology	Population curve	0.47
39	2015	Evolution	Struggle of life	0.91
21	2018	Evolution	Fertile age of the animals (natural selection)	0.66
34	2017			0.50
15	2015			0.78
24	2018	Evolution	Endosymbiosis	0.64
38	2017			0.60
42	2018	Evolution	Mammal's circulatory system	0.51
37	2017	Evolution	Natural selection	0.50
20	2015	Metabolism	CO_2 production	0.44