National level test in science in Latvia for assessing how students explain phenomena scientifically

Pavels Pestovs, Dace Namsone
University of Latvia, Riga, Latvia

Abstract

For successful curriculum implementation in the framework of the project "Competency-based Education Curriculum Development and Implementation" funded from the European Social Fund (ESF), it is essential to develop validated and reliable national level tests with an objective to monitor students' progress. The purpose of the research is to analyse students current situation in explaining natural phenomena scientifically in order to give a clear insight into the assessment process of students skills. At a first glance, a substantial amount (more than 50 %) of 15-16 years old students are capable of explaining natural phenomena in science, however, detailed analysis reveals two problems: both deep and surface student explanations are scored in the same way in test and item assessing criteria are not allowing teachers to assess student skills at different cognitive levels.
Keywords: *explaining phenomena scientifically, natural phenomena in science, skill assessment.*

Introduction

In the framework of PISA (OECD, 2016b) scientific literacy is defined from the point of view of three competencies: explaining phenomena scientifically, evaluating and designing scientific inquiry and interpreting data and evidence scientifically. One of the project's "Competency-based Education Curriculum Development and Implementation" goals is to develop 20 diagnostic tests in order to diagnose students' skill in different ages, 50 % from all diagnostic assessment must be acquired in science, technology, engineering, and mathematics (STEM) area. It is out of the most importance to analyse previous national level tests in order to acquire the insight into the diagnostic system in Latvia. There is a huge gap between national mean percentage in scientific diagnostic assessment of the 15-16 years old students and in OECD PISA results, it is crucial to understand the reason to this difference.

Continuous assessment of students is essential for planning better strategies in student teaching and learning practices. In Latvia, there have been used different assessment strategies, and there are hardly founded evidences that the assessment is shifted only for "teaching for a test". However National Centre for Education has information about variable marking of diagnostic tests, which have been assessed in school by teachers, which reduces the usefulness and reliability of the information (OECD, 2016c). Comparing the results of the national diagnostic tests and PISA nationals results in Latvia reveals controversial information at first glance, but in-depth analysis reveals the difference of methodology and possible explanation about the information and interpretation of it (OECD, 2016a).

The main difference in student skills who are acquired 5th and 6th level proficiency scale level in science, is capability of using abstract scientific ideas or concepts to explain unfamiliar and more complex phenomena, students can offer explanatory hypothesis of novel scientific phenomena from a range of scientific ideas (OECD, 2016b).

The purpose of the study is to find out experience in assessing scientific skill to explain natural phenomena scientifically in Latvia and to analyse the risks and restrictions in today's system in order to plan changes in system of national level test.

This study addresses the following research question:

1. What results students demonstrate in test items which are related to explaining scientific phenomena?
2. What information about students' skill to explain phenomena scientifically is given from national level test to improve testing system?

Methodology of research

National assessment diagnostic test 2017 results for the 15-16 years old students have been analysed in this research. Student national test papers are checked by teachers in schools and scores are delivered for the National Centre for Education of the Republic of Latvia, using electronic system.
The test was completed by 15403 students, with the 35 test elements and maximum score 35 points. Items have been analysed using Classic Test Theory (CTT) and Item Response Theory (IRT) Rasch model.
Correct percentage, discrimination index, percentage endorsing high and low performance have been estimated using ITEMAN program for the CTT. Difficulty parameter (eta) with SE with confident interval (CI) 0.95 has been estimated for each item using freeware R program for the IRT Rasch analysis.
The interest of this study were items with negative IRT Rasch analysis difficulty parameter, in which student ability or attainment level is higher than the item difficulty and where students are explaining natural phenomena scientifically. From the range of items, seven the easiest items were selected, which are related to explaining natural phenomena scientifically. Items which demand students' skill of explaining natural phenomena scientifically and different level of cognitive demand, using both SOLO taxonomy criteria (Hook & Mills, 2012) and Analytical Framework of PISA assessment (OECD, 2016b) were chosen and characterized.
In depth have been analysed 230 papers from 8 schools, both teacher marking scheme and students answers depth were analysed.

Results of Research

Mean score in national diagnostic science test 2017 is 16.7 points with standard error (SE) 5.4. About 25 % of items according to IRT Rasch analysis student ability is higher than the item difficulty. Item characterisation according CTT and IRT Rasch analysis are compiled in Table 1. From the Rasch analysis item-person plot is revealed not enough resolution to the group of students with low and high performances.
It is very useful to get to the point at assessing scientific literacy providing students with tasks with different cognitive demand: low, medium and high.

Table 1. **Characterisation of the test elements of the National assessment**

	SOLO taxonomy	Cognitive demand PISA	Percentage correct, %	Discrimination index	Percentage endorsing low performers	Percentage endorsing high performers	Difficulty parameter	Difficulty parameter with SE
1	Multistructural	Medium	75	0.41	51	92	-1.649	0.019
2	Multistructural	Low	65	0.46	39	84	-1.121	0.017
3	Multistructural	Low	65	0.46	39	85	-1.116	0.017
4	Relational	Medium	63	0.37	43	80	-1.046	0.017
5	Relational	Medium	57	0.52	30	82	-0.773	0.017
6	Unistructural	Low	54	0.39	34	73	-0.627	0.017
7	Relational	Medium	54	0.45	30	75	-0.609	0.017

In-depth analysis of student answers, reveals that a certain percent of answers are not checked correctly by teachers, mostly answers in explanation are unistructural or multistructural according to SOLO taxonomy. It is a rather general practice to score with full credit answers, only if one word hardly matches the explanation. In such practice, we are not capable of evaluation student skill reliably. In-depth analysis of 230 student papers, using SOLO taxonomy, reveals that less than 10 % of students were able to answer the questions using two and more science concepts, demonstrating relational and extended abstract

generalisations. For example, test item, when students are asked to analyse how is it possible to accelerate oxygen production in green plants, using bulb lamp. Two students get maximum point scored for the answers: placing green plant under the Sun and changing the power of bulb lamp and time during the experiment. In future researches it is important to develop reliable rubrics in skill assessment to use them both: in classroom practices and in the national test level.

Conclusions and implications

In the future for the successful developing and implementation new national curriculum, it is highly important, to drive changes which are based in evidence. Longitudinal research, which allow monitoring student progress, using data from validated and reliable diagnostic test system is priority in Latvia. It is highly important to develop diagnostic system, not only in the area of content knowledge, but also in measuring skill development, which requires gathering additional evidences and information. Hattie's research conclusions, about weak skill transferring from different areas, is strongly related to the practices in classroom, in order for students having opportunities in training in different areas.

In-depth student paper study reveals, that only few student demonstrate formulating arguments from different conceptual perspectives, demonstration high level of proficiency scale for science according to PISA framework, therefore using only information about high correct percentage and difficulty parameter are rather misleading.

For future research possible solution in order to develop national level test system is to introduce and adopt electronic testing system in order to use authentic student papers answers and solutions, develop criteria for selection representative student range.

In research literature, surprisingly small is known about how skills are delivered in classroom and how these skills have been assessed in the classroom. For example problem solving skills are not so good correlated, as it would expected, with mathematics and science competency (Csapo & Funke, 2017). In future, developing national level test for skill assessment, essential component is collaborating with psychology experts and teachers in order to validate items, cognitive demand and correlate everyday teaching learning practices in classroom.

References

Biggs, J. B. & Collis, K. F. (1982). *Evaluation the quality of learning: the SOLO taxonomy (structure of the observed learning outcome)*. Academic Press.

Csapo, B., Funke, J. (ed.) (2017). *The Nature of Problem Solving: Using Research to Inspire 21st Century Learning*. Paris: OECD Publishing, http://dx.doi.org/ 10.1787/9789264273955-en.

Hattie, J., Yates, G. C. R. (2014). *Visible Learning and the Science of How We Learn*. New York: Routledge.

Hook P., Mills J. (2012). *SOLO Taxonomy: A Guide for Schools Book 2: Planning for differentiation.*, New Zealand: Essential Resources Educational Publishers Limited.

OECD, (2016a). *PISA 2015 Results (Volume I): Excellence and Equity in Education*. Paris: OECD Publishing, http://dx.doi.org/ 10.1787/9789264266490-en.

OECD, (2016b). *PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic and Financial Literacy*. Paris: OECD Publishing. http://dx.doi.org/10.1787/9789264255425-en

OECD, (2016c). *Education in Latvia, Reviews of National Policies for Education*, OECD Publishing Paris, http://dx.doi.org/ 10.1787/9789264250628-en.

Petty, G. (2014). *Evidence Based Teaching: A Practical Approach.* Nelson Thornes.