DESIGNING & PILOTING ONLINE TESTS AS PART OF A TEACHER COMPETENCE ASSESSMENT

Anete Butkēviča
University of Latvia, Latvia
Inese Dudareva
University of Latvia, Latvia
Dace Namsone
University of Latvia, Latvia
Uldis Zandbergs
Baltic Computer Academy, Latvia
Līga Čakāne
University of Latvia, Latvia
Dace Bērtule
University of Latvia, Latvia

Abstract. *Approaches for competence assessment are becoming increasingly important to plan school development goals, especially during complex educational changes. This paper outlines authors' experience and recommendations from the designing and piloting of teacher online tests as part of a wider competence management process. A developed theoretical model of a universal teacher competence profile is presented. Tests represent four thematic parts related to teacher performance in classroom to facilitate: student cognitive activation, student self-regulation, student collaboration, and leveraging digital. Test questions were developed based on a framework of teacher performance assessment that consists of a structured set of performance level descriptors to help determine teacher level of competence (0-4) according to criteria developed. Online tests were completed by 197 teachers. Test results were analysed to determine their validity. Proposed online tests can be used for interpreting assessment results of teacher knowledge and beliefs necessary for teaching 21st century skills and plan teacher professional development for goal attainment in schools.*
Keywords: teacher competence assessment, test, 21st century skills.

Introduction

Latvia is implementing a compulsory education curriculum reform (Namsone, 2018), and changing the approach in learning process for students to acquire 21st century skills, which is a widespread educational change (Care, Griffin, & Wilson, 2017). High autonomy of Latvian schools (OECD, 2016), as it is in other countries (Mourshed, Chijioke, & Barber, 2010), imply that not all processes in school level can be controlled from above. This emphasise the most
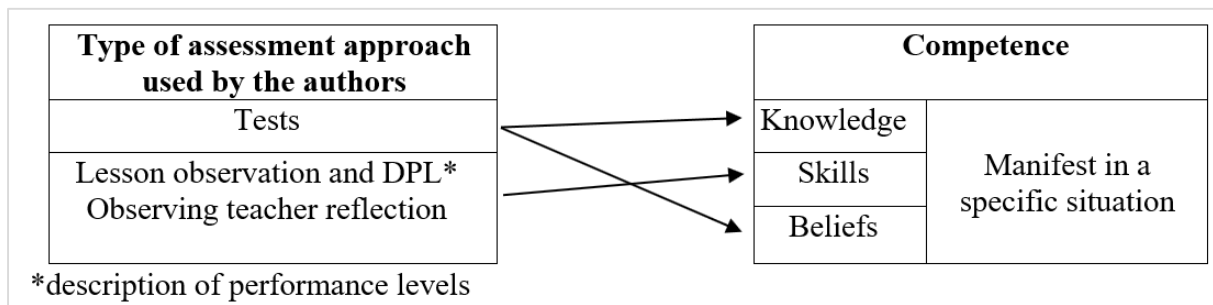
important resource schools as contemporary organizations have – human resources, that is, teachers and their professional competence. Introducing 21st century skills education implies adjustments to how teacher classroom instruction and professional development (PD) should happen.

Currently in Latvia, legitimate Cabinet Regulation prescribes three quality levels in accordance with the list of criteria for assessing performance of teachers mainly on a scale "yes/no" which is done by the school leader (The Republic of Latvia Cabinet Regulation No. 501, 2017). The follow-up of teacher work quality is not based on objective research findings and does not comply with the best practices of other countries. The process of PD is not directly linked with the professional performance of teachers. One of the OECD recommendations for the education system of Latvia state: "Develop a coherent assessment and evaluation framework for informing policy and educational practise" (OECD, 2016).

This paper suggests how evidence-based practices (Mourshed, Chijioke, & Barber, 2010) and structured in-service teacher performance measures through competence assessment may help teachers and school leaders determine competencies available and make development plans for acquiring necessary competencies for achieving new goals that 21st century education brings. In further chapters, the developed construct of teacher competence profile for 21st century teaching and learning will be outlined, first piloting of an online test for assessing parts of teacher competence will be described. First insights of this experience will be presented, further research implications will be discussed.

Background

Our proposed teacher competence consists of the integrated set of knowledge, skills and beliefs that manifest in a specific work situation (Kunter et al., 2013). Teacher competence assessment may be done through using different sources of evidence and different types of assessment approaches (lesson observation, reflective interviews etc.). Each approach focuses on assessing different parts of competence, for example, tests aim to assess teachers' knowledge and decision-making processes (Roelofs & Sanders, 2007; Goodman, Arbona, & Dominguez de Rameriz, 2008). Our proposed tests assess teachers' knowledge and beliefs in an integrated way by including question elements asking teachers' choice of their most typical classroom behaviour. Therefore, the tests reflect only one part of teacher competence, and does not include the skill component (see figure 1). Teacher tests aiming to directly assess pedagogical knowledge are still lacking (Voss, Kunter, & Baumert, 2011).

Type of assessment approach used by the authors	Competence	
Tests	Knowledge	Manifest in a specific situation
Lesson observation and DPL* Observing teacher reflection	Skills	
	Beliefs	

*description of performance levels

Figure 1 ***Types of competence assessment in relation to competence structure based on authors' current experience*** *(Bērtule et al., 2019)*

Competence assessment is part of a wider management and development process in organizations. It may help employees see their performance as linked to achievement of the organization's goals (Campion et al., 2011). Goal setting and goal cascading is part of performance management process, it describes what results an organization is expecting from the employee (Zandbergs et al., 2018). Previous research shows a missing link between goal setting and competence assessment in schools and other organizations in Latvia (Butkēviča et al., 2018).

Objective approaches such as employee competence and performance tests have been used in private sector organizations. It helps to integrate and align human resource systems to enhance and accelerate skill development of employees (Campion et al., 2011). Private sector organizations use computerized tests as an internationally recognized approach by human resource experts and managers for assessment of competences. Tests show that automated tools may be used for objective assessment of competence without overloading of human resource management specialists (Judrups, Zandbergs, & Kazakovs, 2015).

Development and validation of various methods for teacher competence assessment that would correspond to the goal cascading process, competence-based human resource management in formal education providers - schools is a work in progress. For us, competence assessment consists of defining a competence framework, teacher competence profile for assuring objective guidelines of what the competence groups consist of and how they are linked to goals, and through gathering data for assessment in lesson observations (manual) and online tests (automated, information technology (IT) based). These parts of the process would give clear information about decision making in schools. This research describes the development and piloting of teacher in-service tests which is a novel practice in Latvia. In some countries, teachers are subject to testing to assess their general and specialized competencies but it is for granting an entry into the profession. Rarely the results of these tests are used for in-service teacher development (OECD, 2013).

Taking into account the new prioritized changes in Latvian education system, and according to our research project stage, we developed a theoretical framework of four teacher competence groups for teaching 21st century skills. Teaching of 21st century skills is planned to be implemented and started in the education system with the school year 2020-2021. This determines the topicality of the research for schools to set new goals in relation to these changes.

A previously developed category-criteria framework (see table 1) created by the authors (Bērtule et al., 2019) is proposed to be used both for teacher performance assessment (through assessing skills) and for competence profile development, and lastly, for creating tests for assessing knowledge and beliefs. The tests are proposed to be as an alternative and minimize the subjectivity and inaccuracy of an external assessor's activities in assessing teacher knowledge and beliefs - two components of teacher competence.

Questions of the tests were developed based on a framework of teacher performance assessment that consists of a structured set of performance level descriptors to help determine teacher level of performance (score expressed on a scale of 0-4) according to selected criteria for teaching performance to develop 21st century skills (Bērtule et al., 2018). The category-criteria framework consists of eight categories and 13 criteria, each having a qualitative description identifying the performance level (expert, proficient, developing, beginner and not observed). Performance level descriptors where validated based on 145 lesson observations, done by six experts from Interdisciplinary Center for Educational Innovation at the University of Latvia (ICEI UL).

*Table 1 **Selected category-criteria framework for teaching performance to develop 21st century skills** (Bērtule et al., 2019)*

Categories		II 1 Planning	II 2 Teaching	II 3 Classroom environment
IA 1	Student self-regulation	1.1. Learning goals	1.2. Metacognitive skills	
IA 2	Student cognitive activation	2.1. Learning tasks for cognitive depth	2.2.Classroom discourse	
IA 3	Student collaboration	3.1.Learning tasks for collaboration	3.2. Student collaboration	
IA 4	Leveraging digital	4.1. ICT tools	4.2. Meaningful ICT usage	
IB 5 IB 6	Teacher techniques, basic skills	5.1. Lesson design	5.2. Teaching techniques	5.3. Differentiation, personalization, support
		6.1. Curriculum	6.2. Feedback to students	

Previous experience of Baltic Computer Academy (BCA internal working documents such as universal competence profile, tests and competence rubric of instructors conducting PD in various organizations) was analysed and discussed in order to create teacher competence profiles and to conduct competence testing in organizations (Butkēviča et al., 2018). If a school sets a goal for students to develop one of the 21st century skills, then it would be valuable to determine if teachers will be able to reach this goal and if not, then what are the teacher PD needs to do it.

Research questions:
1) What is the theoretical construct (model) for tests to assess knowledge and beliefs necessary for teaching 21st century skills accordingly to potential school goals?
2) How to design the tests that would assess teacher knowledge and beliefs for teaching 21st century skills?
3) What does the test validation process show and what are further improvements needed?

Methodology

The theoretical construct (model) for the tests to assess knowledge and beliefs necessary for teaching of 21st century skills accordingly to potential school goals was created by six experts from ICEI UL. Experts, accordingly, to previously developed category-criteria framework (see table 1) for teaching performance to develop 21st century skills (Bērtule et al., 2019) analysed relevant literature, reviewed development goals formulated by the schools involved in the study and obtained data from the validation process of the mentioned category-criteria framework. Expert discussions were organized to review the theoretical construct (model) involving experts from ICEI UL and the BCA.

Development of tests continued with defining the goal of the tests, defining potential users of the tests, then identifying criteria by which teachers will be assessed according to the structured set of performance level descriptors. A universal and concrete competence profile is presented which the tests are aiming to assess. When developing the test, a data base created from lesson observation was used. The database consists of expert-observers' commentary and citations based on the performance level. Performance levels are structured by criteria, defined in levels (0-4).

All together for developing the four tests, 15 experts were involved, and a workshop was organized. Experts formed four groups according to the competence themes of each test. Each group received those parts from previously mentioned database with criteria referring to the specific test: expert-observers' commentary and citations illustrating teachers' classroom activities according to

performance level. The expert groups created sets of questions referring to each test by using described classroom activity examples (cases) for formulating the multiple choice answers. For each test, a question database was created. It included each question, multiple choice answers and additional data on the question. Test questions and each answer was assigned weights, defined tests' initial size and structure of the question base. By using this database, six experts developed a preliminary version for each test.

The validation of the test took place in a sample of schools that were selected for research purposes. Three groups of teachers were involved, altogether 197 teachers completed the test. Validation of the test took place in the period between 23.04.2018 till 7.05.2018. Distribution of teachers completing each test is as follows: student self-regulation (N=197), leveraging digital (N=49), student cognitive activation (N=85), student collaboration (N=47).

For an additional detailed analysis of data obtained in the validation process SPSS (Statistical Package for the Social Sciences) was used. Descriptive statistics, scale reliability statistics and item reliability statistics (item-total correlation, inter-item correlations, a reliability coefficient McDonald's omega, Guttman's lambda-2 Reliability tests) and where possible, factor analysis was done.

Results and Discussion

The theoretical construct (model) for performance assessment of teaching 21st century skills is designed as a part of the competence management process (see figure 2). It looks at the elements of identified competence assessment parts according to selected school goals. According to selected categories and previously defined competence structure a teacher universal profile is created, that is suitable for selected categories. Teacher profile is developed according to categories, shown in table 1 - Group IA - 21st century skills selected, IB - teacher techniques, classroom management. II - domains of teaching practice - planning (1), teaching (2) and classroom environment (3).
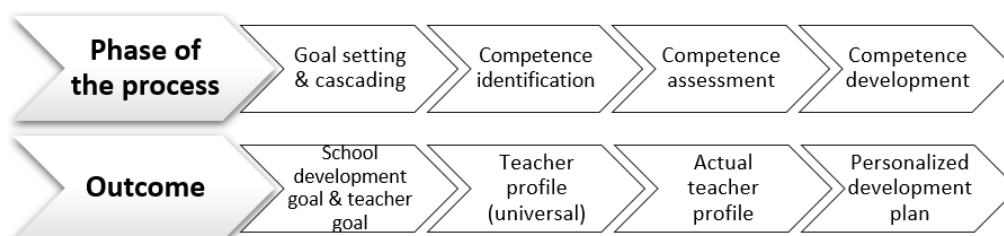

Figure 2 ***Phases of the competence management process and their outcomes***
(ICEI UL archive, 2019)

The preferred performance is aligned with the real performance, obtaining the actual teachers' profile. Actual teacher profile is obtained when the expert-observer determines the performance level according to the developed framework for performance assessment (Bērtule et al., 2019) by lesson observation. Table 2 shows that tests are usable for assessing two elements of competence profile–knowledge and beliefs (see table 2).

Table 2 **Example of match between test questions and category** *(ICEI UL archive, 2019)*

Category	Criteria	Criteria for performance assessment in lesson observation *	Test questions
Student cognitive activation (IA2)	Learning tasks for cognitive depth	2.1.	D_Z_1 D_Z_3 D_Z_4
	Classroom discourse	2.2.	D_Z_2 D_R_1 D_R_2 D_R_3

* see table 1

Development process of the tests include the following stages: I developing questions; II selecting questions and developing tests; III appropriating the tests for online use (such as for MOODLE environment).

In phase I based on the theoretical construct, design of test elements was done. Questions were developed by using previously mentioned category-criteria framework for teaching performance to develop 21st century skills, level descriptors and by selecting specific classroom situations obtained from database of cases gathered in lesson observations (see table 3).

Test question types are: situational analysis tasks, multiple choice questions, matching assignments etc. Question answers are of different type: multiple or single choice; matching judgements or situation descriptions to teachers' characteristic behaviour; yes/no answers; and with different points (from 3 to 7). In phase II questions are ordered into four separate tests, with 29 questions. These tests are initial versions of the tests, that each correspond to chosen category of the developed teacher profile. In phase III test questions are developed in a format accordingly to MOODLE environment.

Data analysis show that the tests offered easily analysable classroom situations. For example, discrimination index in the test "Student self-regulation" is not larger than 0,3 (see figure 4), 50% from the questions in this test the discrimination index is below 0,1 meaning that these questions are invalid. 28% from the questions the discrimination index is between 0,1 and 0,2 meaning that these questions are usable after improvements. This means that the initial version

of the test does not sufficiently discriminate test takers. The difficulty index of questions in all the tests is mostly above 0,5.

Table 3 **Question for assessing knowledge of category "Student cognitive activation"**
(ICEI UL archive, 2019)

| Code	Question type	Question	Assessment	Max. points
D_Z_1	Yes/No	**We want to achieve that students learn to accomplish productive tasks. Please select which statements do you agree or disagree!** **A -** By accomplishing a typical assignment, student learns to use knowledge and skills in new situations, different context. **B** - There should be more assignments in lessons that demand knowledge use in new situations, in other subjects or real life context. **C** – Creating a transfer from known situations to new doesn't require teaching, it develops in long term exercise. **D –** developing use of HOCS can be accomplished through assignment that takes at least half of the lesson and not by using various different, short, unrelated assignments.	1 point for: A – disagree B – agree C – disagree D - agree	4
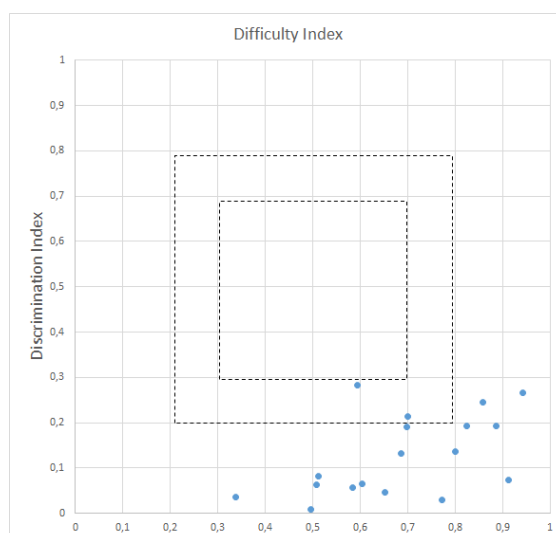
Figure 4 **The relationship between difficulty index and discrimination index of questions of the test "Student self-regulation"** *(ICEI UL archive, 2019)*

For each test question, a violin plot diagram was created (see figure 5 for an example). It shows the frequency of distribution of the data evaluating separately not only each test question but also each test question's choice of answer from the multiple-choice list. This gives detailed information, that will allow for evaluating and further developing each test question and each choice of answer. This is important because there are some valid test questions, but which require adjustments to the offered multiple choice list of answers.

Developed tests will be usable in practice for assessing the selected parts of teacher competence profile, namely, knowledge and beliefs. By improving the tests, it is possible to extend the theoretical construct, adding to the current tests new test questions for personal skills, for example, evaluating reflection skills, thus complementing the inventory of assessment tools.
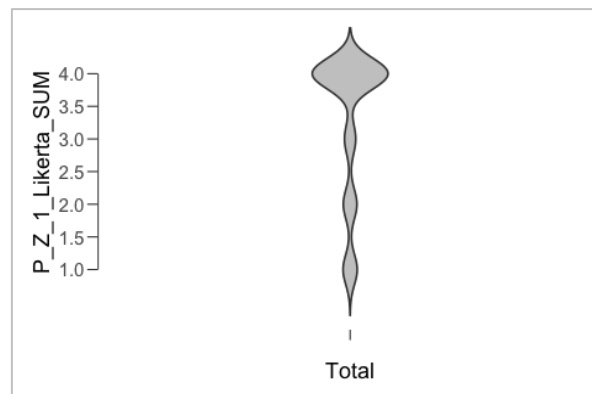
*Figure 5 **A Violin plot diagram for test question P_Z_1** (ICEI UL archive, 2019)*

Using the tests in combination with description of performance levels (DPL), a teacher competence development plan can be drafted prospectively. To conduct research to determine the predictive validity of these tests – do the scores match with results from other sources such as lesson observation and do test scores corelate with student achievement. Previous experience of teacher testing show that there is little empirical evidence about the predictive validity of teachers' performance on the tests as an indicator of classroom effectiveness (Goldhaber & Hansen, 2010).

As in other countries, there is still a need for better teacher tests that would be linked to important teaching knowledge and learning outcomes and validated by independent studies with transparent findings. They would be important accountability mechanisms for states and for programs (Campion et al., 2011).

In further research, it is planned to compare respondents' obtained data from lesson observations with test results. During piloting of next version of the tests, it is necessary to identify reasons that give additional constrains on using the test

and how to increase validity of the test. Further research is also needed to adjust the tests for teacher self-assessment.

When making improvements to the tests, a balance should be found between diversity and the need to attain that the values of question answers differ with the same interval; and that in each test question it is possible to demonstrate knowledge appropriate to criteria in all five levels accordingly to DPL. This would improve the tests' mutual coherence, the number of test questions in tests, the compatibility with parts of the competence profile etc. Invalid test questions have to be sorted out, the test should be complemented with more complex classroom situations. Real life classroom situations gathered from lesson observations were used in the tests. Because in majority of the situations (Bērtule et al., 2019) were on levels 1-2 according to the framework, the data from the tests show that simple classroom situations are being analysed, often in the same performance level.

Conclusions

Four initial test versions for assessing teacher knowledge and beliefs as part of a wider competence construct are developed and first validation in practice is completed. Teacher competence assessment is viewed as a part of competence management process in the context where schools set goals taking into account the planned curriculum reform in the country. The theoretical basis of development of the tests is built on the chooses category-criteria framework by developing teacher competence profiles for teaching 21st century skills. This confirms the developed tests topicality and practical applicability. For the Latvian education context, the tests for in-service teachers can be regarded as innovative. Development of tests' next versions has begun by improving the tests and preparing them for further validation by taking a greater sample of test takers.

Acknowledgments

References

Bērtule, D., Dudareva, I., Namsone, D., Čakāne, L., & Butkēviča, A. (2019). Framework of Teacher Performance Assessment to Support Teaching 21st Century Skills. *13th annual International Technology, Education and Development Conference INTED2019*, Valencia, Spain, 11.-13.03.2019.

Butkēviča, A., Zandbergs, U., Namsone, D., & Briķe, S. (2018). Exploring the Input of Competence Assessment to Goal-setting in Various Types of Organizations. *Society.*

Integration. Education. Proceedings of the International Scientific Conference, Volume VI, 130-141.

Campion, M.A., Fink, A.A., Ruggeberg, B.J., Carr, L., Phillips, G.M., & Odman, R.B. (2011). Doing competencies well: Best practices in competency modeling. *Personnel Psychology*, *64*(1), 225-262.

Care, E., Griffin, P., & Wilson, M. (Eds.). (2017). *Assessment and teaching of 21st century skills: research and applications*. Springer.

Goodman, G., Arbona, C., & Dominguez de Rameriz, R. (2008). High-stakes, minimum-competency exams: How competent are they for evaluating teacher competence? *Journal of Teacher Education*, *59*(1), 24-39.

Goldhaber, D., & Hansen, M. (2010). Race, gender, and teacher testing: How informative a tool is teacher licensure testing? *American Educational Research Journal*, *47*(1), 218-251.

Judrups, J., Zandbergs, U., & Kazakovs, M. (2015). Competence based human resource development solution. *Engineering for Rural Development*, *14*, 669-674.

Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2013). Professional competence of teachers: Effects on instructional quality and student development. *Journal of Educational Psychology*, *105*(3), 805.

Mourshed, M., Chijioke, C., & Barber, M. (2010). *How the world's most improved school systems keep getting better*. McKinsey.

Namsone, D. (Scientific Editor) (2018). *Learning for Competence* (Mācīšanās lietpratībai). Riga: University of Latvia Academic Publishing. DOI: https://doi.org/10.22364/ml.2018

OECD. (2013). *Teachers for the 21st century. Using evaluation to improve teaching*. Paris: OECD Publishing.

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

Roelofs, E., & Sanders, P. (2007). Towards a Framework for Assessing Teacher Competence. *European journal of vocational training*, *40*(1), 123-139.

The Republic of Latvia Cabinet Regulation No. 501 *"Pedagogu profesionālās darbības kvalitātes novērtēšanas organizēšanas kārtība"* (22.08.2017). Retrieved from https://ej.uz/m6o3

Voss, T., Kunter, M., & Baumert, J. (2011). Assessing teacher candidates' general pedagogical/psychological knowledge: Test construction and validation. *Journal of educational psychology*, *103*(4), 952.

Zandbergs, U., Namsone, D., Briķe, S., & Butkēviča, A. (2018). Model of Linking Organization Goals to Employee Competence Management for Formal and Non-formal Education Providers. *13th International Baltic Conference on Databases and Information Systems*, July 1-4, 2018, Trakai, Lithuania.