

countries. However, specific threads are created as well. The Czech Republic follows the strategy not publishing the results from its standardized whole-population testing.

(23) Results from the international comparative assessments PISA and PIRLS/TIMSS were used to explain education system reforms in most of the OECD countries of our interest. These reforms included also the introduction of new evaluation methods (e.g. national testing). On the contrary, the relationship between education performance on one hand and financial allocation on the other was operationalized rather rarely. These facts may be observed also in the case of the Czech Republic.

(24) National testing is the most common method for monitoring students' literacy and numeracy performance level in all the OECD countries of our interest. International comparative assessments PISA and PIRLS/TIMSS are the main source of inspiration for methodological procedures relevant for national tests. Consequently, there are a number of common methodological features in the national testing procedures of the OECD countries of our interest, including test development, test administration, test assessment and test results reporting.

(25) Reading literacy and numeracy are tested in all the OECD countries of our interest. Scientific literacy, social literacy, ICT literacy or foreign language literacy are tested less frequently and in higher grades. However, a higher importance of foreign language literacy in non-English speaking countries is noteworthy.

(26) The choice of a whole-population or sample-based tests is influenced by the main goal of national testing in the OECD countries of our interest. The countries, which intend to evaluate the whole education system and national curriculum, prefer sample-based testing with a longer time period (e.g. Austria, Canada, Ireland, and New Zealand). The countries, which follow the goal to provide information on educational needs of students, prefer annual whole-population testing (e.g. Denmark, Germany, and Slovenia).

(27) There are two strategies of the OECD countries of our interest in their choice of the testing groups of students. Most countries prefer testing of education outcomes at the end of education phases. Thus, national testing verifies whether the education goals and standards were achieved. Some countries (e.g. Ireland, Denmark) test students at the beginning of education phases to recognize their educational needs.

(28) There are only few OECD countries of our interest which use ICT based national testing (e.g. Denmark, Norway). However, the ICT based testing is generally regarded as the future of literacy and numeracy evaluation as this type of testing enables simulations of real world situations.

(29) The quality of education goals and standards influences the quality of national testing in the OECD countries of our interest. Thus, a lower quality of education goals and standards restrains, among others, the opportunity to use national testing for trend evaluations. Education goals and standards serve as benchmarks for trend evaluation. If vaguely formulated, only students of the same studying years are compared.

(30) National testing is a rather new evaluation method in the Czech education system, implemented only in the second decade of the 21<sup>st</sup> century. The testing is focused on reading literacy, on foreign language literacy, and on numeracy in accord with the preceding findings (see point (25)). Results of the national testing are not used for comparative purposes (e.g. school rankings) in the Czech Republic. Thus, the testing serves especially for formative evaluation and for evaluation of the whole education system (see point (26)). There are some opportunities to improve the methodological quality of the national testing in the Czech Republic. These include a broader application of traditional psychometric methods (e.g. Item-