# Comparison of the success rates of candidates for geography studies in solving different types of tasks in 2016 and 2024

Tomáš Matějček[1,2,*], Jakub Jelen[1]

[1] Charles University, Faculty of Science, Department of Social Geography and Regional Development, Czechia
[2] Jan Evangelista Purkyně University in Ústí nad Labem, Faculty of Science, Department of Geography, Czechia
* Corresponding author: tomas.matejcek@natur.cuni.cz

**ABSTRACT**

The aim of this study is to identify changes in geographic literacy over time by comparing the success rates of geography applicants when solving identical tasks used in the geography entrance exam test in 2016 and 2024 at the Faculty of Science of Charles University.

We summarize the results of a comparison of the success rates in selected test tasks and we look at the shift in geographic literacy between two groups of respondents with similar characteristics (age, interest in studying geography at the same university) in different time periods. Through our study, we aim to open up the possibility of using geography entrance exam tests as one of the possible sources for studying the evolution of geographic literacy over time. Longitudinal studies focusing on changes in geographic literacy are still very rare, which we consider to be a research gap. The tasks used in the entrance tests in 2016 and 2024 were compared in order to determine which types of test tasks experienced the greatest change in success rate. The comparison included answers to a total of 25 test tasks, which were intentionally set identically in both years to make such a comparison possible. Answers were available from 269 respondents in 2016 and from 132 respondents in 2024. When evaluating the results, the tasks were divided according to various criteria (thematic focus, category of educational objectives, use of mathematical skills, inclusion of a visual element, etc.). The results indicate a relatively high rate of change in success in solving certain types and groups of test tasks. The results also show changes in the level of geographical literacy of students who come to university from secondary schools. This information could be helpful not only for universities themselves (who will get better information about changes in the level of applicants from secondary school level), but also for secondary school educators and experts engaged in curriculum development (who will get feedback on secondary education results). The results underline the importance of systematically monitoring changes in geographical literacy and call for further research on a larger dataset and across more time points.

**KEYWORDS**

item analysis; entrance exam; geographical education; test task; geographic literacy

Received: 5 May 2025
Accepted: 12 November 2025
Published online: 26 November 2025

Matějček, T., Jelen, J. (2025): Comparison of the success rates of candidates for geography studies in solving different types of tasks in 2016 and 2024. AUC Geographica 60(2), 285–295
https://doi.org/10.14712/23361980.2025.23

## 1. Introduction

The goal of geographical education can be considered, in the most general sense, to be the development of geographic literacy, which we understand as the ability of a person to understand the world from a geographic perspective, to orient themselves in spatial contexts and to understand the relationships between people, places and the environment (for more details, e.g. Binimelis Sebastián et al. 2024; Dikmenli 2014; Soleh et al. 2022; Řezníčková 2003). Geographic literacy thus represents a complex set of knowledge, skills, values, attitudes and competences that enable individuals to understand spatial phenomena, processes and relationships at the local, regional and global levels. It includes the ability to identify and analyze spatial patterns and contexts, interpret geographic information (including maps, data and visualizations) and critically approach their sources and meanings. Geographic literacy therefore includes spatial orientation (knowledge of basic geographical concepts, the ability to read a map, understand where something is and why), understanding natural and social processes, relationship to place (the ability to perceive one's own place in space, to have an overview of one's locality and global context) and the ability to make decisions (to use geographic information in responsible decision-making). The achieved level of geographic literacy at individual levels of school education is commonly verified and assessed in various forms, such as oral exams, essays or written achievement tests), the advantage of which is a high degree of objectivity in evaluating results. The results obtained then indicate the current achieved level in particular aspects of geographic literacy of the group of pupils or students in question. However, research studies that focus on changes in the achieved level of geographical education, or geographic literacy over time, are very rare. Although there are certain partial studies most often focused on specific areas of geographic literacy, longitudinal studies dealing with the change in geographic literacy over time are particularly lacking. We perceive the lack of studies devoted to the development of geographic literacy over time as a research gap. With our study, we would like to open up the possibility of using geography entrance exam tests as one of the possible sources for studying the development of geographic literacy over time.

The aim of this study is to identify changes in geographic literacy over time by comparing the success rates of geography applicants when solving identical tasks used in the geography entrance exam test in 2016 and 2024 at the Faculty of Science of Charles University.

Through the research conducted, we will try to answer the following research question:

How did the success rate of solving different types of test tasks with different thematic focuses change when comparing candidates from different years?

Our research should be understood as a pilot study, using only a limited amount of data. We see its importance mainly in testing the possibilities of evaluating year-on-year changes in the success rate in solving tasks based on various criteria. Nevertheless, we believe that the results of our analysis may be an interesting contribution to the discussion on the development of geographic literacy, which is not yet systematically monitored. The results of our research point to changes in the level of geographic literacy (or more precisely the part of it that can be assessed through entrance tests) over time, for a specific age group, namely secondary school graduates who aspire to study geography at university.

## 2. Geographic literacy and its changes over time

Bendl et al. (2024) state in their study that there are changes in the concept of geographical thinking, a shift from an emphasis on factual knowledge to a focus on geographical thinking and geographical competences. Along with the development of geography, there should also be a change in geographic literacy, but this will certainly not be a parallel process.

The study by Binimelis Sebastián et al. (2024) highlights the low level of geographic literacy of Spanish high school students and states that rote learning (mechanical learning of facts) does not lead to permanent geographic literacy. At the same time, it emphasizes the need for new teaching methods – active, based on working with maps, and developing spatial competences – to increase the geographic literacy of the younger generation.

The importance of strengthening geographic literacy is also mentioned by McFarlane (2024), for example, who emphasizes that geographic literacy is a necessity that has the potential to affect the economy, security, sustainability and cultural understanding.

There are other studies that focus on changes in a specific aspect of geographic literacy. In the USA, for example, research was conducted to map the development of the performance of eighth-grade students in geography between 1994 and 2018 (Solem et al. 2021). This research is based on the National Assessment of Educational Progress (NAEP), a federal program in the USA, sometimes referred to as the "Nation's Report Card". This is a standardized test that monitors the knowledge and skills of American students in various areas – including geography. The results show that American children generally have low levels of geographic literacy, which decreased between 2014 and 2018 (National Center for Education Statistics 2020).

In the framework of scientific literacy research, some geography questions (particularly focused on physical geography) are only occasionally included at a more general level and at the category of pupils in the second stage of primary school (i.e. four or more years younger than the respondents we surveyed) in the regular international surveys TIMSS (Trends in International Mathematics and Science study) and PISA (Programme for International Student Assessment). These surveys deal with the analysis and assessment of pupils' knowledge and skills, and take place at regular intervals. The TIMSS places emphasis on the curriculum and teaching content as well as monitoring the development of knowledge over time (Shi et al. 2016). The PISA survey tests three basic areas: reading literacy, mathematical literacy and science literacy and assesses not only what students know, but also how they are able to work with their knowledge, i.e. the emphasis is on skills, understanding and practical application (Wu 2010).

More detailed analyses of TIMSS and PISA results on science literacy are summarized, for example, by Teig et al. (2022), who identified 82 studies that analyzed data from TIMSS or PISA and focused on different aspects of science teaching and learning. Zhang et al. (2023) developed and validated an instrument for assessing students' science literacy. Kjærnsli and Lie (2004) focused on science literacy in the Nordic countries, analyzing PISA items according to whether they required conceptual understanding or intellectual process skills.

## 3. Admission tests and their evaluation

The admissions process is the gateway to higher education. It is a process during which applicants are assessed according to predetermined criteria and then a decision is made on their admission or non-admission to study. Admissions processes at universities in Czechia are not centrally managed and it is therefore entirely up to each individual university to decide how to organize them. In general, it can be said that the conditions of the admissions process depend primarily on the characteristics of the field of study and the traditions of each university.

If a university uses the option of selecting applicants for study on the basis of the results of an entrance examination, there are several options for how to implement it. Of the various forms (essays, written works, portfolios or oral interviews) tests are the most widespread, usually with automated evaluation using automatic optical scanners (optical mark recognition). Such tests are usually composed of closed questions with one or more correct answers (Pérez-Benedito et al. 2014).

A candidate's performance can be assessed either in absolute terms (reaching a certain threshold required for admission) or on the basis of relative performance (percentile score). In this case, the candidate is compared to other candidates, and the result is expressed relative to other test takers (Ørberg 2018; Frey and Detterman 2004).

Each of the aforementioned admissions options has its advantages and disadvantages, and there are studies devoted to choosing the appropriate tool to determine the qualities of candidates. Researchers also agree that while standardized admissions test scores usually measure cognitive abilities, successful study requires a number of cognitive and non-cognitive attributes that are not verified by admissions tests (Camara and Kimmel 2005; Krumrei-Mancuso et al. 2013; Robbins et al. 2004; Weissberg and Owen 2005).

Silva et al. (2020) states that the best way to select suitable students for university studies is to compare their results in a specific test, which provides objective data, ensuring greater transparency and objectivity of the admission process. On the other hand, he does not consider it appropriate to decide on the admission of students based on their secondary school grades, due to the diversity of schools and the difficulty of comparing academic results of applicants. Other authors oppose the suitability of a combination of different approaches. Bartáková et al. (2018) clearly show that the use of both entrance tests and secondary school grades is justified in university admissions processes for predicting whether a student will be successful in their studies.

Achievement tests are commonly used in entrance exams for the study of geography in Czechia and elsewhere. Tests that can be considered high-quality should meet basic parameters, specifically they should be reasonably difficult and sufficiently sensitive as well as valid, reliable and objective. Objectivity should be ensured by the test being administered to all participants in the same way and by having clear and precisely defined criteria for evaluating the results. The evaluation is therefore independent of the evaluator.

Test reliability indicates how reliable the test is, or rather how stable the results are. An important factor is the consistency of the results in repeated measurements. If a test has high reliability, it will show similar results in different repetitions and under different conditions. Low reliability can be caused by the properties of the test itself, e.g. unclear formulation of the task or inadequate level of tested knowledge related to the tested group (Schindler et al. 2006).

The validity of a test denotes whether the test detects and measures what it is supposed to detect, and that it verifies the tested knowledge at the appropriate level. Validity can be ensured primarily by the relevance of the test's content, i.e. that the test includes all essential parts of the curriculum that the student should know. Pilot testing in a smaller group is also appropriate before implementing the test in practice (Průcha 2009; Štuka and Vejražka 2022; Viktorová and Charvát 2014).

Sensitivity is the ability of the test to correctly identify individuals who have the desired characteristic or feature, e.g. students who have the required skills or knowledge. In the context of achievement tests, sensitivity refers to the ability of the test to distinguish between different levels of performance in students. If the test is highly sensitive, it will be able to distinguish high-performing students from low-performing ones. If students who have better overall knowledge solve a task with great success, while students with worse overall knowledge achieve poor results, this task has high sensitivity (Chráska 1999). A sensitive task, like a sensitive achievement test, is intended to favor students with better knowledge or skills.

There are several ways to determine sensitivity. It is possible to analyze the success rates of individual tasks in a test to determine how well each test task discriminates between high and low-performing students. This is done using the so-called discrimination index, while for a deeper analysis of the sensitivity of a test, the Rasch model, for example, is used (Schindler et al. 2006; Kalhous and Obst 2009; Průcha 2009 and others).

Research on university achievement tests was conducted by, for example, Brožová and Rydval (2014), who analyzed exam tests in mathematics over a period of 13 years. The authors examined the tendencies towards worsening results, the difficulty of the tests and the suitability of the assessment system. The work assessed the quality of the tests using the difficulty index, discriminative ability and reliability.

Entrance exam tests have been analyzed in some previous studies. An analysis of entrance exam tests in mathematics was conducted by Zhang (2023), the properties of tasks in chemistry were analyzed by Šrámek and Teplá (2021); Šrámek and Teplá (2022).

## 4. Materials and methods

For the analysis of test tasks of entrance exams for geography study programs of the Faculty of Science of Charles University, entrance exams from 2016 and 2024 were selected. In 2016, a total of 269 candidates took the test. In 2024, 132 candidates took it. Both entrance exams contained a total of 40 multiple choice test tasks with four answer options, only one answer being correct. Candidates had 60 minutes to complete the tests. Answers were recorded on a separate recording sheet, which was automatically processed and evaluated using computer software.

For both tests, reliability was determined as part of their evaluation by calculating Cronbach's alpha value, which is an internal consistency coefficient that expresses the extent to which individual items of a test or questionnaire measure the same construct and therefore how reliable the test is overall. For the 2016 test, the Cronbach's alpha value was determined to be 0.775 and for the 2024 test it was 0.757. Both of these values indicate good internal consistency of the test, generally considered acceptable for research purposes (Cortina 1993; Tavakol and Dennick 2011). The content validity of the test was assessed by two independent researchers. The main criteria were the thematic focus of the questions and the relevant choice of possible answers.

When selecting test tasks suitable for repetition in 2024, tasks with non-functional distractors, tasks that were too easy (success rate higher than 90%) and tasks that used data that was already outdated in 2024, were eliminated. In this way, a total of 25 test tasks were selected that were used in the same form in the tests in both 2016 and 2024.

These test tasks were subjected to a deeper analysis, which mainly included the calculation of the sensitivity expressed by the discrimination index (Schindler et al. 2006). If a test task has a high discrimination index value, it was solved mainly by students who were successful in the entire test, i.e. these are tasks that favor successful students. Low discrimination index values can also be achieved by tasks that are formulated in a complicated way or tasks for which there are different solution strategies. More successful students may thus try to apply complex solutions, while less successful students only try to guess the correct answer. This discrimination index is calculated using the ULI (upper-lower index) method.

It was used, for example, by Logayah et al. (2024) when evaluating the results of the Geographic Olympiad in Indonesia. An example of its use in other fields is the research of Lucky et al. (2025). In our analysis, we calculated the ULI for individual tasks as a check that there were no test tasks among the analyzed sample that could be ambiguous or incorrectly formulated.

The formula for calculating the discrimination index according to Schindler et al. (2006) is:

$$D = \frac{n_L - n_H}{0.5 \times N},$$

where $n_L$ is number of students from the better half who solved the given task correctly,

$n_H$ is number of students from the worse half who solved the given task correctly,

*N* is total number of students who solved the task.

For the purposes of further evaluation, individual test tasks were classified based on thematic focus, on the educational objectives they verify, as well as on the use of visual elements and the use of mathematical skills.

In terms of thematic focus, the test tasks were divided into five categories: planetary geography, physical geography, social geography (including demography), regional geography (including topography) and cartography. This classification is based on the traditional dichotomy between physical and human (social) geography, as presented by, for example, Cloke et al. (2005) or Hampl (2000).

The evaluation of the educational objectives was carried out using a simplified version of the revised

Bloom's taxonomy, hereinafter referred to as RBT (Amer 2006; Anderson and Krathwohl 2001; Byčkovský and Kotásek 2004). This classification represents the most commonly used classification of educational objectives and is widely used both in theoretical studies and in educational practice. The cognitive process dimension was classified into three categories: Remember, Understand and Apply, with the first two categories corresponding to RBT, and the Apply category also including test tasks that fulfill educational objectives classified in RBT as Analyze and Evaluate. We combined the three categories of Apply, Analyze and Evaluate because the boundary between these them is not always sharp and distinguishing educational objectives between these categories is often difficult (e.g. Amer 2006). The Create category was not included at all, because multiple choice test tasks do not allow the use of this category in practice. Our resulting division of cognitive process dimensions largely (in rough outline) corresponds to the categories proposed by Nimierko (1979). We adopted the assessment of the knowledge dimension from the RBT without modification, thus distinguishing the categories of factual, conceptual, procedural and metacognitive knowledge.

We applied the use of visual elements and use of mathematical skills as sorting aspects because these task characteristics can represent a potential obstacle for a certain type of candidate (especially use of mathematical skills) or, conversely, make solving the task easier for some (use of visual elements). A similar division of tasks was used by Bláha et al. (2024).

The classification of the test tasks (division into categories of thematic focus, prevailing goal according to RBT, knowledge dimension according to RBT, use of visual elements, use of mathematical skills) was done independently by two researchers. The intercoder reliability was 99.2% (that is, 124 out of 125 total items). Only for one task did the evaluators suggest a different classification, which was subsequently discussed.

Tab. 1 shows the resulting classification of individual test tasks. Based on this classification, the tasks

**Tab. 1** Classification of test tasks based on various criteria and their sorting into groups.

| Test task | Thematic focus | Prevailing goal according to RBT | Knowledge dimension according to RBT | Use of visual element | Use of mathematical skills |
|---|---|---|---|---|---|
| 1 | regional geography | remember | factual | no | no |
| 2 | cartography | apply | procedural | no | yes |
| 3 | regional geography | remember | factual | no | no |
| 4 | regional geography | remember | factual | no | no |
| 5 | planetary geography | apply | conceptual | no | no |
| 6 | physical geography | remember | factual | no | no |
| 7 | cartography | apply | conceptual | map | no |
| 8 | regional geography | remember | factual | map | no |
| 9 | regional geography | apply | factual | map and elevation profile | no |
| 10 | cartography | apply | procedural | map | yes |
| 11 | cartography | apply | conceptual | map and elevation profile | no |
| 12 | cartography | apply | conceptual | map | no |
| 13 | regional geography | remember | factual | no | no |
| 14 | planetary geography | apply | procedural | no | yes |
| 15 | planetary geography | apply | procedural | no | yes |
| 16 | physical geography | remember | factual | no | no |
| 17 | physical geography | apply | conceptual | climate diagram | no |
| 18 | physical geography | understand | factual | no | no |
| 19 | physical geography | understand | conceptual | no | no |
| 20 | social geography | remember | factual | no | no |
| 21 | social geography | understand | conceptual | diagram | no |
| 22 | social geography | understand | conceptual | table | no |
| 23 | regional geography | remember | factual | no | no |
| 24 | cartography | apply | conceptual | no | no |
| 25 | physical geography | apply | conceptual | diagram | no |

were divided into groups, for which a summary evaluation was subsequently performed.

The table shows that in terms of thematic focus, the test tasks covered all thematic areas as well as all categories of prevailing goals according to RBT distinguished within the analysis (the Apply category was the most represented). The metacognitive category is missing in the knowledge dimension according to the RBT category, which is not too surprising, since the verification of this dimension would be very unusual in an admission test. Visual elements were represented in ten test tasks, and four test tasks required the use of mathematical skills.

Furthermore, the success rate in % for the years 2016 and 2024 and the difference in the success rate of the given test task between these years were calculated for each test task. This difference was expressed in percentage points with a positive (increase in

**Tab. 2** Task principle and selected indicators of the analyzed test tasks.

| Test task | Task principle | ULI 2016) | Success rate (2016) | Success rate (2024) | Change of the success rate (2016–2024) |
|---|---|---|---|---|---|
| 1 | recognize the relative position of two cities | 0.312 | 56.5% | 55.3% | −1.2% |
| 2 | identify the distance between two points on a map | 0.305 | 57.3% | 61.4% | 4.1% |
| 3 | classify topographic objects by regions | 0.327 | 46.1% | 37.9% | −8.2% |
| 4 | identify false information (about Canada) | 0.156 | 87.0% | 80.3% | −6.7% |
| 5 | identify true information (about Earth's movements) | 0.290 | 57.3% | 45.5% | −11.8% |
| 6 | identify the true definition of the term (pampas) | 0.253 | 82.9% | 75.0% | −7.9% |
| 7 | recognize the area of greatest distortion on the map | 0.223 | 82.2% | 90.9% | 8.8% |
| 8 | select the true names of objects displayed on a blank map (of Poland) | 0.297 | 50.9% | 43.9% | −7.0% |
| 9 | select a transect on a map that corresponds to the displayed elevation profile (North America) | 0.253 | 67.3% | 68.9% | 1.7% |
| 10 | recognize the difference in elevation between two points on a topographic map (imaginary area) | 0.409 | 64.7% | 62.1% | −2.6% |
| 11 | select a transect on a map that corresponds to the displayed elevation profile (imaginary area) | 0.335 | 59.9% | 65.9% | 6.1% |
| 12 | decide which of the drawn watercourses is real (imaginary area) | 0.275 | 56.9% | 56.1% | −0.8% |
| 13 | select the characteristic that corresponds to both given countries (Norway and Chile) | 0.260 | 75.8% | 62.1% | −13.7% |
| 14 | select the local time difference between two locations (defined by geographical coordinates) | 0.223 | 69.5% | 78.0% | 8.5% |
| 15 | select the season during which the Sun is at a given height above the horizon on a given parallel | 0.230 | 48.0% | 31.1% | −16.9% |
| 16 | identify true information (about exogenous processes) | 0.320 | 59.9% | 52.3% | −7.6% |
| 17 | identify the factor responsible for the difference in climatic conditions | 0.283 | 49.4% | 41.7% | −7.8% |
| 18 | decide what the difference is between two terms | 0.245 | 78.1% | 75.0% | −3.1% |
| 19 | recognize which factor is responsible for the difference in salinity and how salinity is affected by it. | 0.223 | 65.8% | 58.3% | −7.5% |
| 20 | choose a pair of languages that are most easily understood by each other | 0.372 | 38.7% | 34.1% | −4.6% |
| 21 | select the process name represented by the diagram (suburbanization) | 0.216 | 75.5% | 82.6% | 7.1% |
| 22 | select a description for the data in the table (imaginary data) | 0.208 | 74.7% | 69.7% | −5.0% |
| 23 | select the relevant area according to the description (protected landscape area) | 0.327 | 42.8% | 40.2% | −2.6% |
| 24 | determine the relative position of two places (defined by geographical coordinates) | 0.372 | 67.7% | 71.2% | 3.6% |
| 25 | select a diagram that corresponds to reality (stream in a river) | 0.141 | 71.0% | 68.2% | −2.8% |
| | Average | 0.274 | 63.4% | 60.3% | −3.1% |

success rate) or negative (decrease in success rate) value. Subsequently, the difference in success rate was calculated for each group of test tasks, which were divided according to the above criteria.

## 5. Results

The research results show that the overall success rate of candidates in solving tasks was lower by 3.1% after eight years. However, visible differences can be identified not only between individual tasks, but also when dividing these tasks into groups according to thematic focus, educational objectives, use of visual elements or use of mathematical skills. While some groups of tasks recorded a visible decrease, only minimal changes or even improvements were visible in other groups of tasks. Detected changes in the success of candidates in solving various types of tasks can also be understood as changes in the geographic literacy of a selected group of respondents of the tested sample.

### 5.1 Results for individual tasks

The basic characteristics of the individual test tasks and their brief description are shown in Tab. 2. The ULI values for each task are also listed. The analysis of 25 test tasks that were repeated in 2024 found that the ULI of most of these selected test tasks was higher than 0.2 (the exception was only two test tasks with ULI = 0.156 and 0.141). It can therefore be stated that test tasks with a relatively high value of the ULI discrimination index were selected for comparison, i.e., test tasks of relatively high quality, with a high level of sensitivity.

Based on the success rate values, most tasks can be rated as moderately difficult, average or moderately easy. Only a few tasks had a success rate lower than 40%, in 2016 the lowest success rate was 38.7% (task no. 20), in 2024 the lowest success rate was 31.1% (task no. 15). We can therefore say that the test did not contain extremely difficult tasks. Similarly, very easy tasks, with a success rate higher than 80%, were represented only exceptionally. The highest success rate in 2016 was 82.9% (task no. 6), in 2024 it was 90.2% (task no. 7). The largest decrease in the success rate in the monitored period occurred in task no. 15 (−16.9%), while the largest improvement occurred in task no. 7 (+8.8%).

### 5.2 Results for specified task groups

As mentioned above, the results of candidates who took the entrance exams in 2024 were 3.1% worse overall than those who took the entrance exams in 2016. However, a summary evaluation of the different test task types shows that some groups experienced a more visible deterioration, while others experienced only slight changes and in one case even improved results. The results of the summary assessment according to various groups of test tasks is shown in Tab. 3.

The greatest decline in success occurred in test tasks thematically focused on physical geography and planetary geography. In physical geography tasks, there was a decrease in success rates for all six test tasks that were included in this category. In the case of planetary geography, while one of the test task saw

**Tab. 3** Success rate of different groups of test tasks and its changes between 2016 and 2024.

| Group of test tasks | | Number of tasks in the group | Average success rate (2016) | Average success rate (2024) | Average change of the success rate (2016–2024) |
|---|---|---|---|---|---|
| thematic focus | planetary geography | n = 3 | 58.2% | 51.5% | −6.7% |
| | physical geography | n = 6 | 67.8% | 61.7% | −6.1% |
| | social geography | n = 3 | 62.9% | 62.1% | −0.8% |
| | regional geography | n = 7 | 60.9% | 55.5% | −5.4% |
| | cartography | n = 6 | 64.7% | 67.9% | 3.2% |
| prevailing goal according to RBT | remember | n = 9 | 60.1% | 53.5% | −6.6% |
| | understand | n = 4 | 73.5% | 71.4% | −2.1% |
| | apply | n = 12 | 63.1% | 61.8% | −1.3% |
| knowledge dimension according to RBT | factual | n = 11 | 62.4% | 56.8% | −5.5% |
| | conceptual | n = 10 | 66.0% | 65.0% | −1.0% |
| | procedural | n = 5 | 59.9% | 58.1% | −1.7% |
| other characteristics | with visual element | n = 10 | 65.2% | 65.0% | −0.2% |
| | without visual element | n = 15 | 62.2% | 57.2% | −5.0% |
| | with mathematical skills | n = 4 | 59.9% | 58.1% | −1.7% |
| | without mathematical skills | n = 21 | 64.1% | 60.7% | −3.4% |

an increase in success rates (a task focused on calculations related to time zones), two tasks saw a visible deterioration in in results – for task no. 15 it was 16.9%, which was the largest decline of all the monitored tasks (this question focused on determining the height of the Sun above the horizon).

Conversely, the only category in which there was an increase in success was the test tasks on cartography. In this category, there was an improvement in results in five out of six test tasks evaluated.

In terms of the predominant objectives according to the RBT, the greatest decrease in success occurred in test tasks that follow objectives in the Remember category and in tasks that verify the factual knowledge dimension. The decrease in success was relatively small in tasks that follow objectives in the Understand and Apply categories and in tasks that verify the conceptual and procedural knowledge dimensions.

Visible differences were noted when test tasks were divided according to whether they contained visual elements and whether mathematical skills were required to solve them. Tasks that did not contain visual elements showed a noticeable decrease in success, while tasks that contained visual elements showed only a minimal decrease in success. Less considerable differences were seen when test tasks were divided according to whether their solution required mathematical skills, but even here differences were noticeable, with a smaller decrease in success for tasks that required some mathematical skills.

## 6. Discussion

This research provides some results worthy of discussion. We consider an important finding to be a visible decrease in success for tasks focused on physical geography and planetary geography and, conversely, the relative increase in success for tasks focused on cartography.

However, it is important to note that thematic focus is not independent of the knowledge dimension according to RBT: in our dataset, tasks classified as physical or planetary geography are predominantly located within the factual knowledge dimension, whereas cartographic tasks more frequently involve procedural or conceptual knowledge. This overlap reflects the broader curricular shift described by Bendl et al. (2024), who point out that the concept of geographical thinking is shifting from the dominance of factual knowledge towards the development of geographical competences, which also requires a gradual redefinition of geographic literacy. Therefore, the differences in success rates across thematic areas may partly reflect these underlying differences in cognitive demands rather than purely thematic characteristics.

In the context of the ongoing reform of the Czech education system, which, among other things, aims to increase the importance of higher cognitive goals compared to simply memorizing facts (for more details, e.g. Trahorsch and Korvasová 2023), the findings of our analysis in the area of educational goals according to RBT can be seen as encouraging, which shows that tasks verifying higher cognitive goals according to RBT showed a smaller decrease in success compared to tasks focused on memorization and the factual knowledge dimension. Similar findings also result from research conducted by Bláha et al. (2024), who analyzed the exam results of university students in the subject of cartography and found that students achieve the worst results in tasks testing factual knowledge, while achieving the best results in procedural knowledge.

On an international scale, Binimelis Sebastián et al. (2024) document that mechanical learning in Spanish schools does not lead to sustainable geographic literacy, and call for approaches that are more active based on working with maps and strengthening spatial competencies, which also corresponds to the intention of the Czech reform.

One may speculate whether the minimal decrease in success rates for tasks with visual elements compared to the more substantial decrease in success rates for tasks without visual elements is related to the higher level of orientation of the younger generation towards audiovisual elements (for more details, e.g. Ateiku et al. 2023; Setyani et al. 2021). Similarly, the causes of the lower decrease in success rates for tasks requiring mathematical skills compared to tasks that do not require mathematical skills may be various, and drawing more serious conclusions would require the analysis of a larger number of tasks over a longer time period. Bláha et al. (2024) found somewhat different results in their research in this case – according to their findings, students achieved worse results in tasks requiring more complex mathematical operations. They also found that the inclusion of visual elements had minimal impact on the success of tasks, but in oral exams, students preferred questions where visualization can be applied or where visual elements are present.

A more detailed analysis would also be needed to examine the potential connection between the increase in success on tasks focused on cartography and the relatively lower decrease in success on tasks with visual elements and on tasks requiring mathematical skills. Cartography tasks usually contained a visual element and solving some of them also required the use of mathematical skills. However, to examine whether these characteristics are related, a larger research sample should be analyzed.

The importance of such analyses is demonstrated by longitudinal studies abroad, such as the NAEP research in the USA (Solem et al. 2021; National Center for Education Statistics 2020). These findings underline that the challenge of strengthening geographic literacy is not unique to the Czech context but represents a broader international trend.

### 6.1 Limitations

We are aware that our study has many limitations and its results should be understood as such. The main limitation of the research is the low number of evaluated test tasks (n = 25). For this reason, the presented study should be understood as a pilot study, the purpose of which was mainly to examine the possibilities of evaluating year-on-year changes in success according to various criteria. The results cannot therefore be considered statistically significant and no fundamental conclusions can be drawn from them. Nevertheless, it can be stated that even the results of the analysis of such a limited sample of examined tasks are stimulating.

Another limitation of the research results from the fact that in the monitored period there was a certain expansion of the group of applicants who were exempted from the entrance exam due to participation in higher levels of subject knowledge competitions (e.g. participation in the national level of the Geography Olympiad, the Secondary School Research Project Competition and other competitive activities). Another reason is the possibility of waiving the entrance exam based on the grade average in selected subjects in secondary school. The percentage of students admitted in this way varies annually, but the value is around 10%.

Certain limitations may also result from the characteristics of the test used. Due to their limited number, the test tasks do not cover the entire breadth of the given disciplines and are not thematically balanced (the proportional representation according to thematic focus is not the same). The same applies to the representation of different types of questions according to educational objectives. There is also a lack of variability in the different types of questions: all questions are multiple choice. However, this results from the nature of the test, because the admission test must be clearly specified and clearly evaluable.

The establishment of categories for classifying test tasks according to educational objectives and the actual classification of tasks into these categories could also be discussed. We are aware that the classification of tasks into individual categories can be subjective, despite the above-described procedure, the aim of which was to minimize the degree of subjectivity. It would also be useful to try out other taxonomies of educational objectives (e. g. the very detailed classification of Tollingerová 1971) for the evaluation of multiple-choice test tasks and to determine the most suitable ones for this purpose. However, this would require a more comprehensive analysis, which is beyond the scope of our study. We consider this to be an interesting challenge and topic for further research.

## 7. Conclusions

The results of our study showed that between 2016 and 2024, when the same test tasks were administered to candidates applying to study geography at the Faculty of Science at Charles University, there was a certain decrease in success rates. However, this decrease was manifested to varying degrees in different types of test tasks, specifically when classifying tasks into different categories according to thematic focus, the educational objectives pursued according to the RBT, as well as other criteria.

While there was a considerable decrease in success rates for test tasks focused on physical geography and planetary geography, there was a relative increase in success rates for tasks focused on cartography. There was a greater decrease in success rates for tasks that tested rote memorization and the factual knowledge dimension, compared to tasks that tested the achievement of higher goals according to the RBT. This finding may indicate a certain shift in the structure of geographic literacy, where tasks based on rote memorization of facts are becoming more difficult for current applicants, while tasks requiring conceptual understanding and application skills are relatively less affected. This trend corresponds to the broader discussion on the transition from factual knowledge towards geographical competences. There was a relatively lower decrease in success rates for tasks that contained visual elements and for tasks whose solution required mathematical skills than for tasks that did not contain visual elements and whose solution did not require the application of mathematical skills. This suggests that candidates may be more accustomed to working with visual and quantitative representations, possibly reflecting their broader experience with audiovisual materials, digital technologies and data-based school assignments.

The results presented here should be viewed with some caution, as they are based on a comparison of only two years of candidates. The uniqueness of our study, however, lies in the fact that the test tasks assessed were identical in both years. Nevertheless, the analysis shows that monitoring year-on-year changes in success rates can provide valuable feedback not only for the refinement of entrance tests (e.g. balancing task types, cognitive levels or thematic coverage) but also for secondary school education or curriculum development.

This pilot study showed that year-on-year comparison of candidates' success rates has considerable potential, and the analysis of a larger volume of data could in the future provide inspiring insights for various readers. Future research should therefore aim to include a larger sample of tasks and more time points, which would allow for more reliable identification of long-term trends in the evolution of geographic literacy among Czech students.

## References


Amer, A. (2006): Reflections on Bloom's Revised Taxonomy. Electronic Journal of Research in Educational Psychology 4(1), 213–230.

Anderson, L. W., Krathwohl, D. R. (Eds.) (2001): A Taxonomy for Learning, Teaching and Assessing. New York: Longman.

Ateiku, J. N., Segbefia, S. K., Gyampoh, A. O. (2023): Effect of Audio-visual Aids on Students Understanding: A Comparative Analysis of JHS Students in the Ho Municipality. Integrated Journal for Research in Arts and Humanities 3(5), https://doi.org/10.55544/ijrah.3.5.4.

Bartáková, I., Chvál, M., Martinková, P. (2018): Predikční validita zkoušek zakončujících středoškolské vzdělávání a přijímacích testů na vysoké školy. Pedagogika 68(1), 66–85, https://doi.org/10.14712/23362189.2017.1016.

Bendl, T., Krajňáková, L., Marada, M., Řezníčková, D. (2024): Geographical thinking in geography education: a systematic review. International Research in Geographical and Environmental Education, 1–19, https://doi.org/10.1080/10382046.2024.2354097.

Binimelis Sebastián, J., Gómez Gonçalves, A., Gómez Trigueros, I. M., Muntaner Guasp, J. J. (2024): Geographic literacy in third-year compulsory secondary education students. Journal of Baltic Science Education 23(6), 1119–1133, https://doi.org/10.33225/jbse/24.23.1119.

Bláha, J. D., Trahorsch, P., Bartůněk, M., Hladík, P. (2024): Critical issues in undergraduate cartographic education: Analysis of final tests and oral examinations. AUC Geographica 60(1), 61–74, https://doi.org/10.14712/23361980.2024.19.

Brožová, H., Rydval, J. (2014): Analysis of exam results of the Subject. Applied Mathematics for IT. Journal on Efficiency and Responsibility in Education and Science 7(3–4), 59–65, https://doi.org/10.7160/eriesj.2014.070303.

Byčkovský, P., Kotásek, J. (2004): Nová teorie klasifikování kognitivních cílů ve vzdělávání: revize Bloomovy taxonomie. Pedagogika 54(3), 227–242 (in Czech).

Camara, W., Kimmel, E. W. (Eds.) (2005): Choosing Students: Higher Education Admissions Tools for the 21st Century (1st ed.). Routledge, https://doi.org/10.4324/9781410612533.

Cortina, J. M. (1993): What is coefficient alpha? An examination of theory and applications. Journal of Applied Psychology 78(1), 98–104, https://doi.org/10.1037/0021-9010.78.1.98.

Cloke, P., Crang, P., Goodwin, M. (2005): Introducing Human Geographies. London: Hodder Education, https://doi.org/10.4135/9781446221235.

Dikmenli, Y. (2014): Geographic literacy perception scale (GLPS) validity and reliability study. Mevlana International Journal of Education 4(1), 1–15, http://dx.doi.org/10.13054/mije.13.43.4.1.

Frey, M. C., Detterman, D. K. (2004): Scholastic Assessment or g? The Relationship Between the Scholastic Assessment Test and General Cognitive Ability. Psychological Science 15(6), 373–378, https://doi.org/10.1111/j.0956-7976.2004.00687.x.

Hampl, M. (2000): Sociální geografie: hledání předmětu studia. Geografický časopis 52(1), 33–40 (in Czech).

Chráska, M. (1999): Didaktické testy: příručka pro učitele a studenty učitelství. Brno: Paido, ISBN 80-85931-68-0 (in Czech).

Kalhous, Z., Obst, O. (2009): Školní didaktika. Praha: Kosmas, ISBN 978-80-7367-571-4 (in Czech).

Kjærnsli, M., Lie, S. (2004): PISA and scientific literacy: Similarities and differences between the Nordic countries. Scandinavian Journal of Educational Research 48(3), 271–286, https://doi.org/10.1080/00313830410001695736.

Krumrei-Mancuso, E. J., Newton, F. B., Kim, E., Wilcox, D. (2013): Psychosocial factors predicting first-year college student success. Journal of College Student Development 54(3), 247–266, https://doi.org/10.1353/csd.2013.0034.

Logayah, D. S., Ruhimat, M., Arrasyid, R., Islamy, M. R. F. (2024): Item analysis of National Geography Olympiad multiplechoice questions (MCQs) in Indonesia. Cogent Social Sciences 10(1): 2354971, https://doi.org/10.1080/23311886.2024.2354971.

Lucky, A., Janštová, V., Novotný, P., Mourek, J. (2025): Quantifying ento-literacy: development and validation of an international insect-focused attitude and knowledge survey instrument. International Journal of STEM Education 12: 11, https://doi.org/10.1186/s40594-025-00532-8.

McFarlane, D. A. (2024): Critical Review of Geographic Illiteracy in the United States: A Multifaceted Analysis. Journal of Geography, Environment and Earth Science International 28(1), 11–25, https://doi.org/10.9734/jgeesi/2024/v28i1741.

National Center for Education Statistics (2020): The Nation's Report Card: 2018 U.S. History, Geography, and Civics at Grade 8 – Highlights Reports (NCES 2020017). U.S. Department of Education. Available online: https://nces.ed.gov/use-work/resource-library/report/statistical-analysis-report/nations-report-card-2018-u-s-history-geography-and-civics-grade-8-highlights-reports/ (accessed on 25 March 2025).

Niemierko, B. (1979): Taksonomie celów wychowania. Kwartalnik pedagogiczny 24(2), 66–67 (in Polish).

Ørberg, J. W. (2018): Uncomfortable encounters between elite and "shadow education" in India – Indian Institutes of Technology and the Joint Entrance Examination coaching industry. Higher Education 76, 129–144, https://doi.org/10.1007/s10734-017-0202-5.

Pérez-Benedito, J. L., Aragón, E. Q., Alriols, J. A., Medic, L. (2014): Optical Mark Recognition in Student Continuous Assessment. IEEE Revista Iberoamericana de Tecnologías del Aprendizaje 9(4), 133–138, https://doi.org/10.1109/RITA.2014.2363005.

Průcha, J. (2009): Pedagogická encyklopedie. Praha: Portál, EAN 9788073675462 (in Czech).

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., Carlstrom, A. (2004): Do Psychosocial and Study Skill Factors Predict College Outcomes? A Meta-Analysis. Psychological Bulletin 130(2), 261–288, https://doi.org/10.1037/0033-2909.130.2.261.

Řezníčková, D. (2003): Geografické dovednosti, jejich specifikace a kategorizace. Geografie 108(2), 146–163, https://doi.org/10.37040/geografie2003108020146 (in Czech).

Schindler, R., Brož, F., Hulešová, M., Chvál, M., Kovařík, Š., Matoušek, J., Mičienka, M., Smetáčková, I. (2006):

Rukověť autora testových úloh. Praha: Centrum pro zjišťování výsledků vzdělávání, ISBN 80-239-7111-5 (in Czech).
Setyani, W. A., Jumadi, J., Darmawan, A. S. (2021): The Implementation of Audio Visual Media in Problem Based Learning Model to Improve the Problem Solving Skills.
Proceedings of the 6th International Seminar on Science Education (ISSE 2020). Advances in Social Science, Education and Humanities Research 541, 563–568, https://doi.org/10.2991/assehr.k.210326.081.
Shi, W.-Z., He, X., Wang, Y., Fan, Z. G., Guo, L. (2016): PISA and TIMSS Science Score, Which Clock is More Accurate to Indicate National Science and Technology Competitiveness? Eurasia Journal of Mathematics, Science and Technology Education 12(4), 965–974, https://doi.org/10.12973/eurasia.2016.1239a.
Soleh, T. A., Ruhimat, M., Yani, A. (2022): Factor analysis: geographic literacy level of high school students in North Bandung area. International Journal of Research and Review 9(12), 519–528, https://doi.org/10.52403/ijrr.20221257.
Solem, M., Vaughan, P., Savage, C., De Nadai, A. S. (2021): Student- and School-Level Predictors of Geography Achievement in the United States, 1994–2018. Journal of Geography 120(6), 201–211, https://doi.org/10.1080/00221341.2021.2000009.
Silva, P., Nunes, L. C., Seabra, C., Balcao Reis, A., Alves, M. (2020): Student selection and performance in higher education: admission exams vs. high school scores. Education Economics, 28(5), 437–454, https://doi.org/10.1080/09645292.2020.1782846.
Šrámek, M., Teplá, M. (2021): The Qualities of Tasks in General Chemistry Based on the Analysis of Entrance Procedure at the Faculty of Science, Charles University. Scientia in Educatione 12(2), 32–40, https://doi.org/10.14712/18047106.2068.
Šrámek, M., Teplá, M. (2022): The qualities of tasks in organic chemistry and biochemistry based on the analysis of entrance procedure at the Faculty of Science, Charles University. Scientia in Educatione 13(2), 51–63, https://doi.org/10.14712/18047106.2219.
Štuka, Č., Vejražka, M. (2022): Testování a hodnocení studentů na VŠ. Praha: Univerzita Karlova, https://doi.org/10.14712/9788024651088 (in Czech).
Tavakol, M., Dennick, R. (2011): Making sense of Cronbach's alpha. International Journal of Medical Education 2, 53–55, https://doi.org/10.5116/ijme.4dfb.8dfd.
Teig, N., Scherer, R., Olsen, R. V. (2022): A systematic review of studies investigating science teaching and learning: over two decades of TIMSS and PISA. International Journal of Science Education 44(12), 2035–2058, https://doi.org/10.1080/09500693.2022.2109075.
Tollingerová, D. (1971): Úvod do teorie a praxe programované výuky a výcviku. Odborná výchova 21(5), 143–146 (in Czech).
Trahorsch, P., Korvasová, V. (2023): Jaké je postavení geografie v probíhající kurikulární reformě? Diskuze klíčových témat geografického vzdělávání v rámci revize Rámcového vzdělávacího programu pro základní vzdělávání. Informace ČGS 42(1), 1–11 (in Czech).
Vašková, M. (2013): Přijímací řízení na vysoké školy. Diplomová práce. Masarykova univerzita, Právnická fakulta, Brno. Available online: https://theses.cz/id/fsffjy/ (accessed on 10 March 2025, in Czech).
Viktorová, L., Charvát, M. (2014): K problematice validity testů studijních předpokladů. Testfórum 3, 3–15, https://doi.org/10.5817/TF2014-3-13 (in Czech).
Weissberg, N. C., Owen, D. R. (2005): Do Psychosocial and Study Skill Factors Predict College Outcomes? Comment on Robbins et al. (2004). Psychological Bulletin 131(3), 407–409, https://doi.org/10.1037/0033-2909.131.3.407.
Wu, M. (2010): Comparing the Similarities and Differences of PISA 2003 and TIMSS. OECD Education Working Papers 32. Paris: OECD Publishing, https://doi.org/10.1787/5km4psnm13nx-en.
Zhang, H. (2023): A Comparative Analysis of the 2018–2021 College Entrance Examination Mathematics National Paper (Science) questions. International Journal of Education and Humanities 8(1), 122–130, https://doi.org/10.54097/ijeh.v8i1.7244.
Zhang, L., Liu, X., Feng, H. (2023): Development and validation of an instrument for assessing scientific literacy from junior to senior high school. Disciplinary and Interdisciplinary Science Education Research 5: 21, https://doi.org/10.1186/s43031-023-00093-2.