Geography tests in the Finnish Matriculation Examination in paper and digital forms – an analysis of questions based on revised Bloom's taxonomy

Eerika Virranmäki[a]*, Kirsi Valta-Hulkkonen[b] and Anne Pellikka[c]

[a]*Geography Research Unit, University of Oulu, Finland

[b]Teacher Training School, University of Oulu, Finland

[c]Faculty of Education, University of Oulu, Finland

*Corresponding author: Eerika Virranmäki, Geography Research Unit, P.O. Box 3000, 90014 University of Oulu, Finland, +35829 448 1713, eerika.virranmaki@oulu.fi

Kirsi Valta-Hulkkonen, Teacher Training School, P.O. Box 9300, 90014 University of Oulu, Finland, +35829 448 3876, kirsi.valta-hulkkonen@oulu.fi

Anne Pellikka, Faculty of Education, P.O. Box 2000, 90014 University of Oulu, Finland, +35829 448 7572, anne.pellikka@oulu.fi

Acknowledgement: The authors thank the two anonymous reviewers for their valuable suggestions and overall comments for improving the quality of the manuscript.

Funding details: This work was supported by the LUMA (STEM) Centre of the University of Oulu.

Disclosure statement: No potential conflict of interest was reported by the authors.

Data availability statement: The data (geography test questions, although only in Finnish) of this study is openly available at the website of the Finnish public service media company, YLE https://yle.fi/aihe/artikkeli/2015/12/15/yo-kokeet-maantiede

Notes on contributors:

Eerika Virranmäki (MSc, Doctoral Researcher) is doing her PhD study in the Geography Research Unit at the University of Oulu. Her research interests are in the field of geography education primarily in the upper secondary schools, geographical knowledge and changing pedagogies in geography education and assessment. ORCid: https://orcid.org/0000-0002-5212-860X

Kirsi Valta-Hulkkonen (PhD) is working as a lecturer in the Teacher Training School in the Faculty of Education at the University of Oulu. Her research interests are in the fields of geography and biology education. ORCid: https://orcid.org/0000-0002-1308-2530

Anne Pellikka (MSc, Doctoral Researcher) is a university teacher in the didactics of biology, geography and health education in the Faculty of Education at the University of Oulu. Her research interests are in the fields of biology, geography and environmental studies in the context of teacher education. ORCid: https://orcid.org/0000-0002-8969-3602

Geography tests in the Finnish Matriculation Examination in paper and digital forms – An analysis of questions based on revised Bloom's taxonomy

Abstract

This study examines the possible change in cognitive processes and geographical knowledge requirements during the digitalisation process of the Finnish Matriculation Examination in Geography (digitalised in the autumn of 2016). The theoretical background is derived from a revised Bloom's taxonomy, which is applied as a framework for the content analysis of the questions (n = 331) of the geography tests between the autumn of 2013 and the spring of 2019. The results show that the questions mainly require an understanding of conceptual and factual knowledge. Because of the digitalisation, questions that require remembering have been reduced, whereas questions that require analysing have been increased. In addition, more comprehensive use and more extensive production of materials are required in the tests. Based on our research findings, we emphasise the need for a careful re-evaluation of the kind of cognitive processes and knowledge that should be assessed in geography tests, both nationally and internationally.

Keywords: revised Bloom's taxonomy; geography education; national matriculation examination; Summative evaluation; Finland

Abbreviations

FNAE Finnish National Agency for Education (former FNBE, Finnish National Board of Education) is national agency which is responsible to develop education and training from early childhood education and care to general and vocational upper secondary education.

USS Upper Secondary School is three-year post-compulsory school which leads to the national Matriculation Examination. Students are usually aged 16–18 and they have studied earlier ten years in Finnish education system: pre-primary education for 6-years-olds (one year) and basic education (comprehensive schools) for children aged 7–16 (nine-years).

M.E. Matriculation Examination is a dominant, national summative large-scale examination taken at the end of the Finnish Upper Secondary school. When students are passing the Examination they are able to continue their studies at universities or universities of Applied Sciences.

FMEB Finnish Matriculation Examination Board is responsible for administering the Matriculation Examination. It takes care of the arrangements and execution of the Examination.

More information about Finnish education system https://www.oph.fi/en/education-system

1. Introduction

1.1. Geography education and assessment

In recent years, research topics in geography education have attempted to highlight its importance in a rapidly changing world. There is interest in examining 'twenty-first century skills' (Béneker, Palings, & Krause, 2015; Pauw, 2015, p. 308) and powerful geographical knowledge (see e.g. Maude, 2015, 2016, 2018; Slater, Graves, & Lambert, 2016; Virranmäki, Valta-Hulkkonen, & Rusanen, 2019). It has also been underlined that the aim of geography education should focus on education for young peoples' capabilities as future global citizens (see e.g. Lambert, Solem, & Tani, 2015; Uhlenwinkel, Béneker, Bladh, Tani, & Lambert, 2017). These ideas include learning skills such as critical thinking, creativity, collaboration, self-management, information literacy, and digital skills (Pauw, 2015, p. 311). Additionally, they include understanding of the knowledge of the world and geographical perspectives, and using scientific concepts and methods, and being able to participate in public debate (Maude, 2018, pp. 3–7; Virranmäki et al., 2019, pp. 11–12) and thinking geographically (Uhlenwinkel et al., 2017, p. 10).

However, it is important to notice that many studies (see e.g. Butt & Lambert, 2014, p. 9; Brooks, Qian, & Salinas-Silva, 2018, p. 8; Uhlenwinkel et al., 2017, p. 10) note that national school systems and educational aims vary widely. Thus, geography as a school subject may be part of the social sciences, the humanities, or the natural sciences. In Finland, geography is linked to biology, chemistry, physics, and mathematics. It is a named subject in the Finnish national secondary curriculum, with one compulsory and three optional courses for all students. Despite the different geographical education settings, Butt and Lambert (2014) note that there seems to be a general 'body of knowledge that is common across the globe' (p. 1). Chang and Seow (2018) also argue that 'there is some congruence in general understandings of what the goals of geographical education might be' (p. 32). Additionally,

the International Charter on Geographical Education (Commission of Geography Education of the International Geographical Union [IGU-CGE], 2016) emphasises the importance of geography as a discipline:

> 'Geography is concerned with human-environment interactions in the context of specific places and locations and with issues that have a strong geographical dimension like natural hazards, climate change, energy supplies, migration, land use, urbanization, poverty and identity. Geography is a bridge between natural and social sciences and encourages the 'holistic' study of such issues' (IGU-CGE, 2016, p. 10).

According to Favier and van der Schee (2012), modern geography education should be seen as 'more like an activity that students can engage in' (p. 666) rather than fixed factual knowledge. Students should be able to acquire and use geographical knowledge and develop geographical skills for doing geography (Favier & van der Schee, 2012, p. 666; van der Schee, Notté, & Zwartjes, 2010, p. 7). It is therefore important that we can help students to learn not only knowledge, but skills and attitudes (see also Chang, & Kidman, 2019, p. 2).

The fact that assessment can improve teaching and learning is often ignored (Stoltman, Lidstone, & Kidman, 2014, p. 193; Wertheim & Edelson, 2013, p. 15). However, assessment has an impact on teaching and learning, including its content and style (Baird, Andrich, Hopfenbeck, & Stobart, 2017, p. 340), and 'the ways students are assessed on their knowledge' (Ormond, 2019, p. 6) play a role in determining what knowledge is taught to students. Thus, 'assessment is a goal-setting activity' (Baird et al., 2017, p. 340). It is important to examine 'the ways in which we assess geographical learning' (Chang & Seow, 2018, p. 38; see also Bijsterbosch, van der Schee, & Kuiper, 2017, p. 19). Although assessments are not always transparent (Ormond, 2019, p. 18), making assessment procedures and processes observable may increase student success (in tests), but it may also lead to 'teaching to test' and a 'standards-based curriculum' (Torrance, 2011, p. 464), in which teaching of standards is seen as crucial for education. These aspects and the research topics

and recognised differences between the educational settings presented above raise questions about the kind of cognitive processes and knowledge dimensions we are assessing in geography.

In their large review article of assessment in geography, Lane and Bourke (2017, p. 12) summarise their findings by pointing out that more research is needed concerning the knowledge and skills students should develop during their geographical education. Furthermore, Wertheim and Edelson (2013) call for thinking about 'what is being assessed and how it is assessed' (p. 15). We also need assessment frameworks that assess competencies (Leat, Thomas, & Reid, 2012, p. 401) and 'challenging and divergent assignments -- encouraging students to search for strategies that fit the task at hand and their own potential' (Hooghuis, van der Schee, van der Velde, Imants, & Volman, 2014, p. 245).

1.2. The digitalisation of the Finnish national matriculation examination

Education and assessment are largely intertwined within national contexts. It is therefore necessary to explain more precisely the reform of the Matriculation Examination carried out in the Finnish assessment scene. The Finnish national Matriculation Examination (M.E.) is the dominant large-scale summative assessment of learning outcomes administered at the end of Finnish upper secondary school (USS) (approximately 35,000 participants per year). It consists of at least four tests in different subjects. The mother tongue exam is the only compulsory subject. The participants select the three other compulsory tests from different subjects, and they can also include optional tests. These tests are held biannually, in the spring and autumn, simultaneously in all Finnish USSs, and they aim to examine whether students have acquired the skills and competences defined in the National Core Curriculum for General Upper Secondary Schools (see the Finnish National Board of Education [FNBE], 2016). According to the test policy, students taking part in the tests are usually aged 18 and have studied for 13 years in the Finnish education system. After successfully completing the

compulsory tests and receiving an upper secondary education or a vocational upper secondary certificate, students receive the matriculation examination certificate.

The major reform was carried out during 2016–2019 when the M.E. was converted to digital format following a decision of the Finnish government in 2011 (Valtioneuvoston kanslia, 2011, p. 33). Geography was one of the first tests to be digitalised, in the autumn of 2016; mathematics was the last, in the spring of 2019. During the digitalisation process, changes were made to the geography test structure. In the paper-based test, there were ten assignments, of which students had to answer six (there were no compulsory assignments, cf. digital tests). The last two assignments (marked with a + sign) were designed to be more demanding in the knowledge and cognitive process categories. The total maximum score was 42 points (Finnish Matriculation Examination Board [FMEB], 2017a). The new digital geography test consists of three different parts I–III (Part I is compulsory for all participants) containing a total of nine assignments, of which students are required to answer five. The maximum score is 120 points, and each assignment can give 20–30 points (FMEB, 2017b).

Changes have also been made to the knowledge and cognitive process demands of the geography test. The FMEB (n.d., p. 2) stated a policy applying revised Bloom's taxonomy (see also Table 1) as a guiding principle for designing the new digital test assignments. This principle is reflected in the different parts of the geography test described by the FMEB (the Geography Subject Section of the FMEB, 2018): Part I is compulsory for all students, containing assignments that evaluate the students' primary knowledge in geography, mainly remembering and understanding; Part II focuses on applying knowledge but overlaps with assignments that require students to analyse information; Part III requires comprehensive knowledge of geography and problem-solving skills in different contexts by requiring students to analyse, evaluate, and create knowledge. Parts II and III both contain four assignments, and students choose two assignments from each part.

The Finnish M.E. will gain significantly in importance in the spring of 2020, because the majority of USS students will be accepted into higher education based on their success in it, and the role of the current entrance exams will be reduced. This, as Torrance (2017) and Lambert (2017) also point out, is a sign that education is increasingly going at the direction in which its purpose is to 'sort and grade young people for the workforce' (Lambert, 2017, p. 16). However, Pellegrino (2017) states that assessment should not merely rank students but should produce 'information about the nature of student understanding' (p. 365).

1.3. Aims and research questions

Research by Lane and Bourke (2017) suggests 'that while there had been a number of changes in educational standards and curriculum documents internationally, assessment practices had not always adjusted to align with these changes' (p. 11). In Finland, we have a unique opportunity to examine assessment in geography, because the M.E. has been digitalised, and the new question types and test structures have been formed by applying revised Bloom's taxonomy. To our knowledge, there is no published research on the reformed geography test questions of the Finnish M.E., and there has been almost no research into revised Bloom's taxonomy in geography assessment and education in practice. Our paper contributes new knowledge to this field.

Our aim is to examine the possible change in the cognitive processes and geographical knowledge requirements of the geography tests during the Finnish M.E. digitalisation process. We provide an analysis of the cognitive processes and geographical knowledge that are actually assessed in two different forms of the tests. We do this by analysing the Finnish M.E. geography tests' questions (digital since the autumn of 2016) between the autumn of 2013 and the spring of 2019 by using revised Blooms' taxonomy as a framework for our analysis. Our research question is: How have the Finnish M.E. geography test questions changed during the digitalisation process, when examining:

1) the geographical cognitive processes and knowledge required;
2) the types of attached material and related cognitive processes and knowledge required;
3) the types of assignment and related cognitive processes and knowledge required.

2. Theoretical background

The theoretical background for this article has been derived from revised Bloom's taxonomy, which we also use as a framework to analyse our extensive data about geography tests in Finland. In the methods section of this article (see Section 3), we explain more accurately how we have interpreted this framework in the context of geography education. We now examine the theoretical aspects.

The use of revised Bloom's taxonomy, named the Taxonomy Table (Table 1) by Anderson and Krathwohl in 2001 (Anderson et al., 2014), is one way to approach assessment. Krathwohl (2002) describes the Taxonomy as 'a framework for classifying statements of what we expect or intend students to learn as a result of instruction' (p. 212). This idea was originally Benjamin S. Bloom's. He believed the Taxonomy was more than a tool for measurement; it could serve as 'common language about learning goals' (Krathwohl, 2002, p. 212). Airasian and Miranda (2002, pp. 253–254) state that the Taxonomy Table can be used to analyse nationwide assessments and to determine the kind of cognitive processes and types on which we should focus, rather than revealing examples of specific test items in the assessments.

Table 1. Revised Bloom's taxonomy, the Taxonomy Table (Anderson et al., 2014).

The knowledge dimension	The cognitive process dimension					
	Remember	Understand	Apply	Analyse	Evaluate	Create
A) Factual						
B) Conceptual						
C) Procedural						
D) Metacognitive						

The taxonomy has two dimensions: cognitive process and knowledge. The cognitive process dimension includes six categories: remembering, understanding, applying, analysing, evaluating, and creating (Anderson et al. 2014, p. 5). The first three can be called lower-order cognitive skills (LOCS); the last three of them can be called higher-order cognitive skills (HOCS). LOCS measure aspects such as the ability to remember, understand, or solve routine problems; HOCS relate to areas such as the ability to select and analyse knowledge, solve real-life problems, and think critically (e.g. Tikkanen & Aksela, 2012; Zoller & Pushkin, 2007). When students are dealing with HOCS-type questions, they cannot rely solely on memory (Anderson et al., 2014, p. 71), and teachers must 'assume a less direct role in facilitating student learning' (Anderson, 2005, p. 110).

The knowledge dimension contains four types of knowledge: factual, conceptual, procedural, and metacognitive knowledge. Concrete factual knowledge is the knowledge, or 'basic elements', students should know to be able to understand or solve problems of a particular discipline. When these 'basic elements' form a functional larger structure, and the interrelationships between the concepts are visible, it is called organised conceptual knowledge. Procedural knowledge refers to knowledge about how to do something, and how to think and inquire within the discipline. Metacognitive knowledge is abstract, and it differs from the other knowledge categories: it is self-knowledge; self-awareness; and 'knowledge of one's own cognition' (Anderson et al., 2014, p. 46). It has been suggested that the focus in assessment should be on HOCS-questions and the metacognitive knowledge dimension (e.g. Airasian & Miranda, 2002; James & Gipps, 1998; Tsaparlis & Zoller, 2003).

2.1. Earlier research findings

A few researches into revised Bloom's taxonomy and geography questions in geography education has been conducted. Lane and Bourke (2017, p. 7) found five papers in their review article, which concerned cognitive demand and knowledge types in geography. Bijsterbosch et al. (2017) used revised Bloom's taxonomy to analyse 'internal school-based geography examinations of the final exam in pre-vocational secondary education in the Netherlands' (p. 17); Wertheim and Edelson (2013) analysed assessment items in 'large-scale standardised assessments and classroom assessments' (p. 18) in K-12 classrooms in the US. They both concluded that the majority of analysed test items focused on the categories of remembering and understanding mainly factual and conceptual knowledge. Bijsterbosch et al. (2017, p. 26) found that evaluating and creating, the most complex cognitive processes, was rarely evaluated. Wertheim and Edelson (2013, p. 18) made the same observation, acknowledging that students' ability to perform evidence-based reasoning was lacking in the analysed tests.

There have been similar findings in research on geography textbook questions (Jo & Bednarz, 2009; Mishra, 2015), although they have used a geospatial thinking taxonomy. Jo and Bednarz (2009, p. 9) revealed that only 13 per cent of the analysed questions related to the evaluation and creation processes, while more than half the questions required the recalling of information. Mishra (2015, p. 123) concluded the same: only 22 per cent of the questions required evaluation and creation, while almost half the questions were related to remembering. However, Yasar (2009) and Yang (2013) examined changes in textbook questions after national educational reforms in China and Turkey, concluding that questions requiring higher-order thinking had become more numerous and better designed.

It has been noted that digital technologies, especially digital representations (such as digital globes and maps) and geographic information systems are suitable for developing students' HOC-skills (see e.g. Collins, 2018; Favier, & van der Schee, 2014; Palladino, &

Goodchild, 1993), and that modern technology offers new possibilities for teaching and learning, as well as testing, by combining 'different sources such as digital maps, photos and video simultaneously or successively' (van der Schee et al., 2010, p. 7). For example, Liu, Bui, Chang, and Lossman (2010) concluded that when taught with digital geographic information systems, students demonstrated more HOC-skills than students in a control group, whereas Favier and van der Schee (2014, p. 227) showed that using different layers of interactive maps could help students to demonstrate HOC-skills better than when using paper maps. However, Collins (2018, p. 139) pointed out that there was no adequate empirical evidence to show that paper maps were worse for improving students' HOC-skills than digital maps. Digital technologies have not only introduced new materials such as digital maps, but it is argued that digitalisation shifted the writing process towards digital writing, in which the planning process is understood to be shorter than when writing with pen and paper (e.g., Nordmark, 2014; Rautapuro, & Harjunen, 2017).

3. Data and methods

The study's data-gathering and analysis process is illustrated in Fig. 1. The research data consists of the Finnish M.E. geography tests between the autumn of 2013 and the spring of 2019. The M.E. was in a paper-based format between the autumn of 2013 and the spring of 2016, and in digital format between the autumn of 2016 and the spring of 2019. Altogether, 12 tests containing 114 assignments with 331 questions (when all the subsections of the assignments were considered) were analysed in two different periods, between November 2017 and January 2018, and again between August 2019 and October 2019. The statistical analysis were conducted in March 2020. The analysed geography tests can be found on the website of the Finnish public service media company (YLE, 2019), as they are public information in Finland.

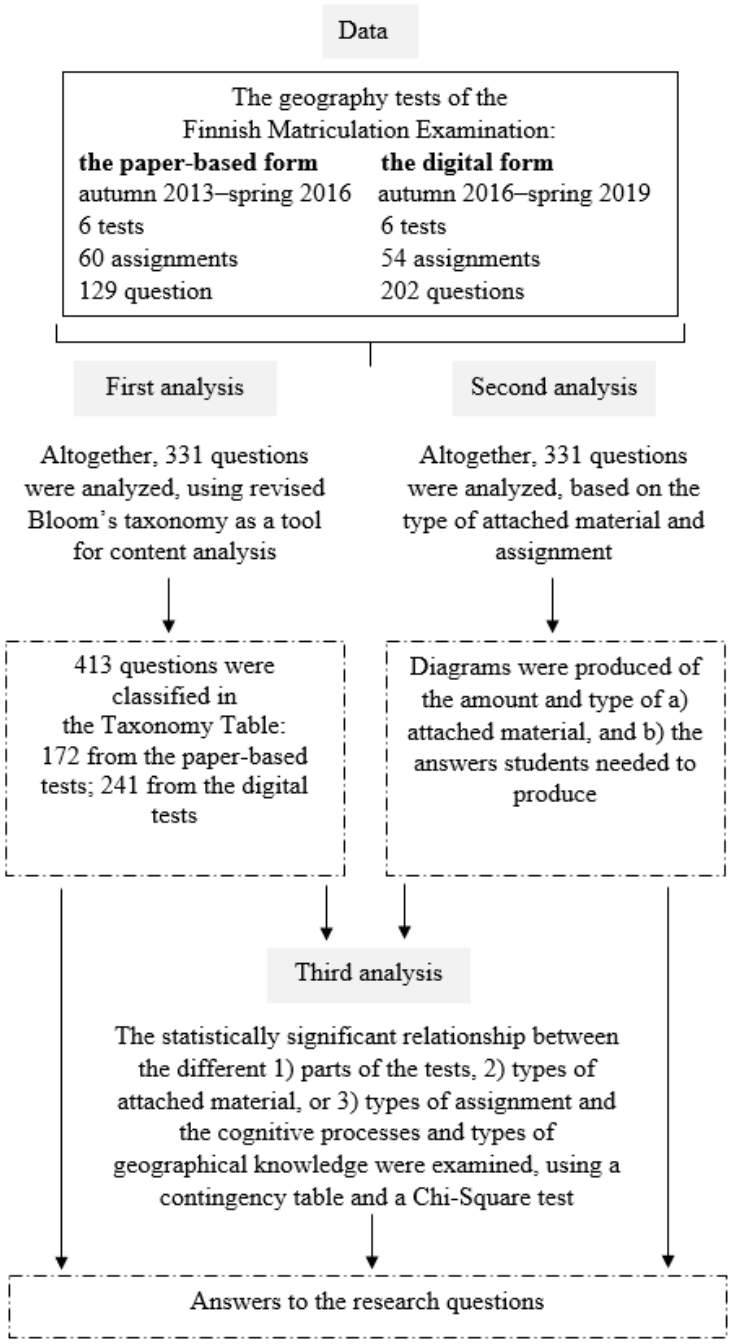

Figure 1. Phases of data gathering and the analysis process in this study.

In our first analysis, we classified questions in accordance with the Taxonomy Table (see Table 1). The study's first author collected the raw data, meaning the questions from the examinations, both from the paper-based and digitalised examinations. She a) read the

questions of both examinations; b) categorised them according to revised Bloom's taxonomy; and c) produced a preliminary written categorisation of the questions. After this procedure, the first author sent the preliminary categorisation to the second and third authors individually. They then read and analysed the preliminary categorisation. The third phase was conducted by all three authors. In this third phase, the categorisation was scrutinised collaboratively, so all three authors could agree on the accuracy of the categorisation. Finally, the first author produced an in-category comparison of the analysed questions. This enabled us to examine and compare the categorised questions within the categories in which they were classified. The reliability of our analysis was further strengthened by dialogue between the three researchers during the article-writing process.

It is important to note that cognitive process categories are hierarchical, but that they also overlap (Krathwohl, 2002, p. 215). Therefore, in our classification, one assignment or question may be divided into one or more categories, based on the kind of cognitive process it requires (see also Bijsterbosch et al., 2017, p. 24; Radmehr & Drake, 2018, p. 46; for a contrasting method see Tikkanen & Aksela, 2012). For example, the question 'Assignment 1. a) Explain what concept X means. Analyse its spatial variation on earth', is grouped in the categories of 'remember' and 'analyse', because it requires two kinds of cognitive process. Consequently, a total of 413 questions are grouped according to the Taxonomy Table.

In our analysis, we have also used the 'Features of a good answer' provided by the FMEB (YLE, 2019). These documents present model answers to every assignment so the criteria for the assessment are visible in practice. This lends reliability to our analysis, enabling us to secure our interpretations of the questions based on the principles which the definers attempted to implement when compiling the questions. Additionally, geography textbooks used in USSs and our own knowledge of geography for interpreting questions were

utilised. We were thus able to decide if the questions were completely new to students (for a contrasting method, see Tikkanen & Aksela, 2012, p. 263).

In our second analysis, we typed 331 questions for different categories based on the type and amount of the attached material related to them and the kind of answer students were required to produce (e.g. a text or diagram). We therefore produced diagrams to represent the variation in attached material and producing-type assignments of the Finnish M.E. geography test questions between the autumn of 2013 and the spring of 2019. Again, the first author produced the preliminary categorisation of the questions and sent it to the second and third authors for their review. The results of this categorisation were produced collaboratively.

In our third analysis, we conducted statistical analysis to examine whether there were any statistically significant relationships between the different 1) parts of the tests (i.e. the last two assignments or Part III –assignments), 2) the types of attached material, or 3) the types of assignment and cognitive process (categorised in two groups, LOCS and HOCS) and types of geographical knowledge. We used a contingency table to analyse our data to determine whether there was any relationship between the variables used and identify the frequency distribution between different variables. We also used a Chi-Square test to analyse whether the contingency between the variables was statistically significant. Again, the raw data for the statistical analysis and the analysis themselves was produced by the first author. The second and third authors then evaluated the results individually. Finally, the decision to include the results of the statistical analysis in this paper was analysed and discussed collaboratively at a joint meeting.

3.1. Using revised Bloom's taxonomy as a framework for content analysis

We used a qualitative approach and theory-driven content analysis method in our first and second analysis of the research data in this study. We applied revised Bloom's taxonomy to the context of geography education and used this framework to analyse the Finnish M.E.

geography tests' questions between the autumn of 2013 and the spring of 2019. Tables 2 and 3 describe our interpretation of the cognitive processes and the knowledge dimensions applied in classifying geography test questions.

	Lower-order cognitive skills (LOCS)			Higher-order cognitive skills (HOCS)		
	Remember	Understand	Apply	Analyse	Evaluate	Create
Definition:	Retrieve relevant knowledge from long-term memory (recognising, recalling)	Construct meaning from instructional messages, including oral, written, and graphic communication (interpreting, exemplifying, classifying, summarising, inferring, comparing, explaining)	Carry out or use a procedure in a given situation (executing, implementing)	Break material into constituent parts and determine how parts relate to one another and to an over-all structure or purpose (differentiating, organising, attributing)	Make judgements based on criteria and standards (checking, critiquing)	Put elements together to form a coherent or functional whole; reorganise elements into a new pattern or structure (generating, planning, producing)
Questions require the student to:	Recognise geographical symbols from the presented material Remember simple facts, and recall concepts and pictures from long-term memory	Describe different geographical phenomena by listing and explaining concepts Give examples, compare and classify geographical concepts Infer and explain how geographical processes work from information presented Translate and summarise the information of a given representation into a different form	Apply simple geographical models or theories to explain different phenomena Apply knowledge of geographical methods, e.g. draw a map from the given material	Select relevant information from the presented material and organise it to form a coherent conclusion so that causalities between phenomena or concepts are visible Analyse the values and attitudes of the presented material	Draw conclusions and judgements from the given phenomena based on known criteria and standards by justifying views Be critical, i.e. critical thinking is visible	Put elements together in a way that it forms a coherent whole that is new way to see phenomena and hypothesise how the phenomena are going to proceed – i.e. by answering a 'what then?' -question Show creative and holistic thinking by reorganising elements
Examples of geography questions from the Finnish M.E. during 2013–2019:	Recognising map symbols, e.g. name the biomes presented on the map Recalling geographical concepts, e.g. what 'renewable energy resources' means Draw a picture of the structure of the atmosphere relying on memory alone	Interpreting the given map to illustrate how the percentage of the population aged between 7 and 15 varies regionally in Finland Giving regional examples for each phenomenon Classifying concepts, e.g. whether they are part of the climate or the weather Summarising the content of a video in an essay title Inferring how a mountain in a given picture was formed Comparing how the climate in the given area differs from the climate in Finland Explaining causes of earthquakes	Executing a simple calculation, e.g. the surface area Implementing, e.g. using information of spatial data and how it can be used to determine the location of a new mall	Differentiating the information on maps and charts about free-time living in Finland Organising, e.g. structuring a view of how global risks affect Finland Attributing, e.g. noticing the values presented in the text (not found in M.E. 2013–2019)	Checking what advantages are related to measuring happiness through self-assessment Critiquing, e.g. judging the social impact of the Olympics by analysing the pros and cons the Olympics have for the organising city and its citizens	Generating e.g. hypothesising how to improve world food production by using given material Planning an outline of an essay based on the given material Producing a plan or a map of a trip to North America

Table 2. The cognitive process dimension of the Taxonomy Table applied in the context of the geography test questions (based on Anderson et al., 2014).

	A) Factual knowledge	B) Conceptual knowledge	C) Procedural Knowledge	D) Metacognitive knowledge
Definition:	Aa) Knowledge of terminology Ab) Knowledge of specific details and elements	Ba) Knowledge of classifications and categories Bb) Knowledge of principles and generalisations Bc) Knowledge of theories, models and structures	Ca) Knowledge of subject-specific skills and algorithms Cb) Knowledge of subject-specific techniques and methods Cc) Knowledge of criteria for determining when to use appropriate procedures	Da) Strategic knowledge Db) Knowledge about cognitive tasks, including appropriate contextual and conditional knowledge Dc) Self-knowledge
Questions require the student to:	Demonstrate knowledge of simple facts or specific details, concepts, elements, or phenomena	Demonstrate knowledge of the causalities between concepts by connecting things Explain theories, models, structures, classifications, categories, principles, and generalisations with the help of examples	Apply knowledge of geographical methods and criteria for using methods in certain situations Demonstrate understanding of the grounds of the specific method and be able to use it in a real situation	Know how to read and memorise geographical texts Know how to construct a geographical answer for different cognitive tasks (e.g. recalling versus critiquing) Display knowledge of one's own strengths and weaknesses
Examples of geography questions from the Finnish M.E. during 2013–2019:	Explain what poverty means or interpret map symbols Give examples of a specific volcano type or hazard occurring in a specific area	Explain how absolute and relative poverty differ Describe where and why geographical phenomena occur Analyse spatial differences and reasons for population growth globally	Calculate distances between places marked on the map Explain how different map types are produced Analyse how mobile applications utilise spatial data Explain what kinds of spatial data you could use when analysing routes	(not found in M.E. 2013–2019) Set goals for reading a geographical text Find the best way to write an analysis of geographical phenomena Find out how well you can solve different cognitive tasks

Table 3. The knowledge dimension of the Taxonomy Table applied in the context of the geography test questions (based on Anderson et al., 2014).

When using the cognitive dimension of the Taxonomy Table in this study (see Table 2), we understand that the difference between remembering and understanding is that remembering concentrates solely on recognising and recalling discrete elements, while understanding relates to constructing connections between prior and new knowledge. However, we accept that remembering knowledge is a prerequisite for other cognitive processes, because such information is used in more complex tasks (see Anderson et al., 2014, p. 66, 70). Understanding is both the largest and a very comprehensive cognitive process, because it includes many subcategories, all of which are an important part of education and learning. These categories have huge potential to enhance the learning of geography (see also Bijsterbosch et al., 2017, p. 18). However, when students understand something, they are mainly using their knowledge by listing things related to a given task and explaining concepts.

Applying is understood in our research as the capability of using procedures to solve familiar or unfamiliar exercises and problems (Anderson et al., 2014, p. 77). Consequently, students can use their prior knowledge in different situations and apply geographical models, theories, or procedures to produce an answer for a given task. Analysing is a continuation of understanding, because it requires students to process and organise knowledge from different sources (Anderson et al., 2014, p. 79). Although not all the questions include attached material, they are still classified in the category of ‘analyse’ if they expect students to form a coherent structure from previously disorganised knowledge. This means that when students analyse something, they recognise causalities and can decide which information is suitable in the given situation. This differentiates it from the process of understanding.

We understand that the process categories of ‘understand’ and ‘analyse’ are interrelated with the more complex categories of ‘evaluate’ and ‘create’. In the process of ‘evaluate’ the student is checking or critiquing something from a certain perspective, and

drawing conclusions and judgements based on that evaluation. It is important to note that not all judgements are understood as evaluative. This means that students should use known criteria and standards (Anderson et al., 2014, p. 83) to construct justifiable arguments and demonstrate critical thinking, and draw a firm conclusion based on the analysis the student has done. The category of 'evaluate' is thus the extension of the category of 'analyse'.

Concerning the category of 'create', we are referring to the idea that when students are creating, they are generating different solutions to a problem, providing a plan to solve it, and actually solving it by producing something (Anderson et al., 2014, p.85). Creating requires students into synthesise scattered material into an organised whole or original product. Anderson et al. (2014, p. 85) state that writing can be seen as producing, i.e. creating, something (see e.g., Tikkanen, & Aksela, 2012). However, it is important to notice that not all writing is on the level of 'create', because creating also requires a deep understanding that goes beyond the students' own experiences and knowledge (Anderson et al., 2014, p. 85). In this sense, creating relates to creative and holistic thinking and synthesis. Creating thus differs from the lower processes of 'understand', 'analyse' and 'evaluate'.

Table 3 shows the knowledge dimension of the Taxonomy Table used in the context of geography education. An important distinction should be made here between factual and conceptual knowledge. In this study, we understand terms, facts, and concepts as factual knowledge if they are merely separate parts of information. If they are connected to form a larger system of ideas or transfer to learners' everyday experiences, they are understood as conceptual knowledge (see Anderson et al., 2014, p. 42). Procedural knowledge is understood as knowledge that is related to the knowledge of how to do something subject- or discipline-specific (e.g. skills, methods, and techniques), whereas metacognitive knowledge is knowledge about the student's own strengths and weaknesses in studying geography. It refers

to the broader skills or strategies of learning and problem solving (see Anderson et al., 2014, pp. 52–60).

4. Results

4.1. Comparing the cognitive and knowledge requirements of the paper-based and digital forms of the tests

Our results show that the paper-based (Table 4) and digital (Table 5) geography test questions (n = 172 and n = 241 respectively) of the Finnish Matriculation Examination between the autumn of 2013 and the spring of 2019 are mainly LOCS-type questions (71 % and 70 % respectively). In both forms of test, the majority of questions are based on the category of 'understand' (34 % in the paper-based tests; 41 % in the digital tests). However, in the paper-based tests, the second cognitive category is remembering (28 %), and the third is analysing (19 %); in the digital tests, the opposite applies (21 % and 23 % respectively). The questions requiring students to analyse have increased during the digitalisation process. In both forms of test, eight per cent of the questions are classified in the category of 'apply', while the two least required categories are 'create' and 'evaluate'. There is a remarkable decrease in the category of 'create', from the paper-based tests' eight per cent to the digital tests' three per cent. In the digital tests, questions from the category of 'evaluate' have increased, although by only one percentage point.

In both forms of test, the majority of the questions emphasise conceptual (49 % of the questions in the paper-based tests and 47 % of the questions in the digital tests) and factual (35 % and 39 % respectively) geographical knowledge. Knowledge of geographical theories, models, and structures (24 % and 28 % respectively) is especially emphasised. There is a minor difference concerning the second knowledge dimension: in the digital tests there are more questions requiring a knowledge of terminology (24 %) than questions requiring a knowledge of specific details and elements (15 %). In the paper-based tests, the opposite is

the case (13 % and 22 % respectively). There is a lack of questions requiring procedural and metacognitive geographical knowledge. Only 16 per cent of the questions in the paper-based tests and 13 per cent of the questions in the digital tests are on the procedural level, whereas none of the questions are on the metacognitive level.

		Cognitive process dimension							
		LOCS (lower-order cognitive skills)			*HOCS (higher-order cognitive skills)*				
Knowledge dimension		*Remember*	*Understand*	*Apply*	*Analyse*	*Evaluate*	*Create*	*Total*	
A) Factual knowledge	Aa) Knowledge of terminology	9(15)	3(6)				1(1)	13(22)	
	Ab) Knowledge of specific details and elements	10(18)	9(15)		2(3)		1(2)	22(38)	35(60)
B) Conceptual knowledge	Ba) Knowledge of classifications and categories		5(9)	1(1)	2(3)			8(13)	
	Bb) Knowledge of principles and generalisations	1(1)	8(13)		5(9)	1(1)	4(7)	18(31)	49(85)
	Bc) Knowledge of theories, models, and structures	1(1)	8(14)	2(4)	10(18)	1(1)	2(3)	24(41)	
C) Procedural knowledge	Ca) Knowledge of subject-specific skills and algorithms	8(13)	1(1)	2(3)			1(1)	10(18)	
	Cb) Knowledge of subject-specific techniques and methods	1(1)	1(1)			1(1)		2(3)	16(27)
	Cc) Knowledge of criteria for determining when to use appropriate procedures			3(6)				3(6)	
D) Metacognitive knowledge	Da) Strategic knowledge								
	Db) Knowledge about cognitive tasks, including appropriate contextual and conditional knowledge								
	Dc) Self-knowledge								
	Total	28(49)	34(59)	8(14)	19(33)	2(3)	8(14)	100(172)	
			71(122)			29(50)			100(172)

Table 4. Distribution (percentage (number)) of the paper-based (autumn 2013 and spring 2016) geography test questions (n=172) in the Finnish Matriculation Examination, according to revised Bloom's taxonomy table.

		Cognitive process dimension							
		LOCS (lower-order cognitive skills)			*HOCS (higher-order cognitive skills)*				
Knowledge dimension		*Remember*	*Understand*	*Apply*	*Analyse*	*Evaluate*	*Create*	Total	
A) Factual knowledge	Aa) Knowledge of terminology	16(39)	8(19)					24(58)	
	Ab) Knowledge of specific details and elements	3(8)	10(25)		2(4)			15(37)	39(95)

B) Conceptual knowledge	Ba) Knowledge of classifications and categories	1(1)	3(7)			1(1)		4(9)	
	Bb) Knowledge of principles and generalisations		10(23)		5(13)	1(2)		16(38)	47(114)
	Bc) Knowledge of theories, models, and structures		8(19)	1(1)	15(37)	1(5)	2(5)	28(67)	
C) Procedural knowledge	Ca) Knowledge of subject-specific skills and algorithms	1(2)		3(8)				4(10)	
	Cb) Knowledge of subject-specific techniques and methods		2(6)	1(1)	1(2)			4(9)	13(32)
	Cc) Knowledge of criteria for determining when to use appropriate procedures			4(10)			1(3)	5(13)	
D) Metacognitive knowledge	Da) Strategic knowledge								
	Db) Knowledge about cognitive tasks, including appropriate contextual and conditional knowledge								
	Dc) Self-knowledge								
	Total	21(50)	41(99)	8(20)	23(56)	3(8)	3(8)	100(241)	
			70(169)			30(72)			100(241)

Table 5. Distribution (percentage (number)) of the digital (autumn 2016 and spring 2019) geography test questions (n=241) in the Finnish Matriculation Examination, according to revised Bloom's taxonomy table.

In examining the structure of the examinations by statistical analysis, we found that the last parts of the tests (the last two assignments from the paper-based tests and the Part III – assignments from the digital tests) were more demanding than the other parts of the test. The results from the paper-based tests showed that of the last two assignments, 60.9 per cent were HOCS-type questions, while 77.6 per cent of the other questions were LOCS-type questions. For the digital tests, we found that the majority of the Part III –questions (51.3 %) were HOCS-type questions, while the majority of the questions from Parts I and II (94.5 % and 75.6 % respectively) were LOCS-type questions. A Chi-Square test was used to analyse whether the differences in the cognitive processes between the last two assignments and other assignments in the paper-based tests, or between Part III and Part I or II in the digital tests,

were statistically significant. At a significance level of $p < 0.05$, the test results were significantly different: $p = 0.000$ in both tests.

The statistical analysis revealed that the last two assignments from the paper-based tests and the Part III –assignments from the digital tests were statistically different (a Chi-Square test, $p = 0.000$ in both forms of test) from the earlier parts of the test with regard to their geographical knowledge requirements. The results showed that the questions from the paper-based tests mainly required the use of conceptual knowledge (75.4 % of the questions from the last two assignments and 44.1 % of the questions from the other assignments) and factual knowledge (20.3 % and 39.2 % respectively). However, there was a difference if the question required the use of procedural geographical knowledge: 4.3 per cent of the questions from the last two assignments and 16.8 per cent of the questions from the other assignments required the use of procedural knowledge. In examining the digital tests, we found that the questions from the Part I assignments mainly required factual knowledge (81.8 %), while the questions from Parts II and III mainly required conceptual knowledge (58.1 % and 63.5 % respectively). Procedural level questions were mostly found in Part III.

4.2. Comparing the types of attached material according to the paper-based and the digital forms of the tests

The number of attached materials included in the geography tests between the autumn of 2013 and the spring of 2019 are presented in Fig. 2. Comparing the paper-based (autumn 2013–spring 2016) and digital (autumn 2016–spring 2019) tests, the number of attached materials has increased since the digital tests started: the average of attached materials is 12 in the paper-based tests and 19.5 in the digital tests. However, the materials provided have decreased during the last two digital tests to almost the same level as the paper-based tests. Maps, texts, diagrams, and drawings have almost always been part of the attached material of

the geography tests during the research period. Along with digitalisation, new materials, including charts, and statistics, videos and interactive maps, have emerged.

There is a statistically significant difference (Chi-Square test, p = 0.006 in the paper-based tests and p = 0.027 in the digital tests) between the types of attached material and the requirement of LOC- or HOC-skills. In the paper-based tests, the questions with drawings or maps as attached material required more LOC-skills (77.8 % and 63 % respectively); the questions with texts or diagrams as attached material required more HOC-skills (63.6 % and 68 % respectively). Similarly, in the digital tests, the questions with texts or diagrams as attached material require more HOC-skills (73.7 % and 62.5 % respectively), and drawings, maps, interactive maps, videos, and charts as attached material require more LOC-skills (64.3 %, 61.4 %, 78.6 %, 62.5 %, and 56.7 % respectively). This indicates that the questions with non-traditional attached material (i.e. the interactive maps and videos that emerged with digitalisation) require mainly LOC-skills. However, there is no statistically significant difference between the types of attached material and the knowledge dimension of the questions.

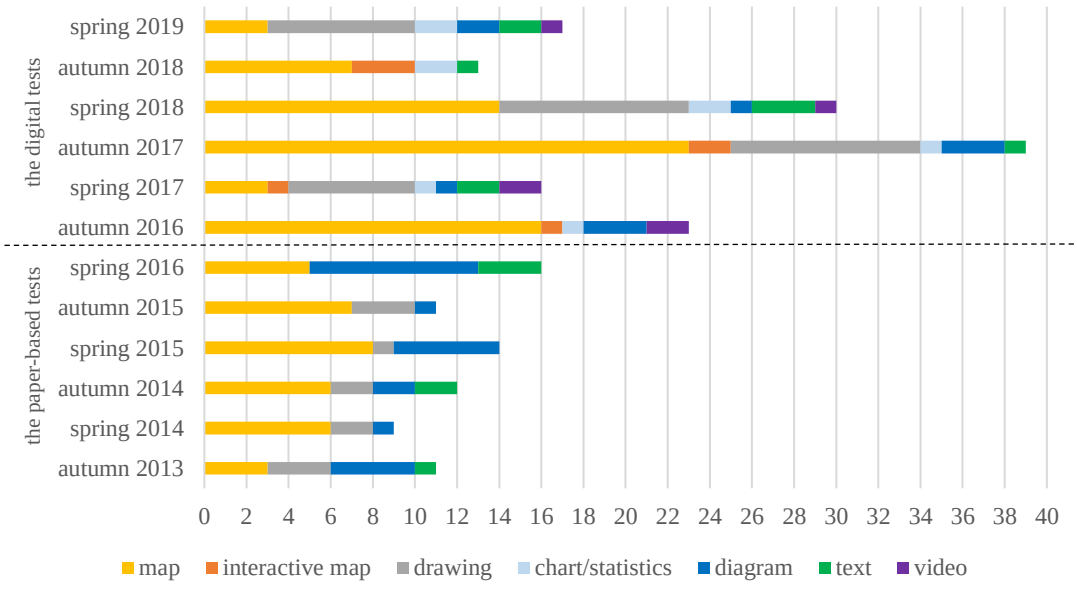

Figure 2. The number and type of the attached material of the Finnish Matriculation Examination geography tests between the autumn of 2013 and the spring of 2019.

4.3. Comparing the types of assignment according to the paper-based and the digital forms of the tests

Fig. 3 presents the number and types of the geography test assignment between the autumn of 2013 and the spring of 2019. Assignments requiring students to produce text-based answers were the basis of geography tests during the research period. In the paper-based tests (between the autumn of 2013 and the spring of 2016), there were questions that required students to produce text, maps, and drawings, and two questions requiring calculation. In the digital tests (between the autumn of 2016 and the spring of 2019), most questions still required text-based answers, but mathematical exercises (requiring statistical numeracy skills) increased, while simple drawings decreased. Totally new elements were multiple-choice questions, and questions requiring the production of diagrams (e.g. climate diagram and altitude profile) or charts.

In the digital tests, there was a statistical (Chi-Square test, $p = 0.000$) difference between the types of assignment and the cognitive requirements of the questions. 43.3 per cent of the questions requiring text-based answers required the use of HOCS-skills; the corresponding value was 1.7 per cent for questions requiring another kind of answer (multiple-choice, drawing, map, diagram, calculation, and chart). This shows that the new types of assignment, multiple-choice questions, diagrams, and charts, mainly required LOC-skills. There was no similar statistically significant difference in the paper-based tests.

In examining the difference between the types of assignment and the knowledge requirements of the questions, we found that the differences were statistically significant in both forms of test (Chi-Square test, $p = 0.000$). In the paper-based tests, the questions requiring text-based answers mainly required conceptual knowledge (59.6 %); while other assignment types mainly required procedural knowledge (84.2 %). Similarly, in the digital

tests, questions requiring text-based answers mainly required conceptual knowledge (64.3 %). However, other types of assignment mainly required factual knowledge (69.5 %). This is explained by the large number of multiple-choice questions in the digital tests.

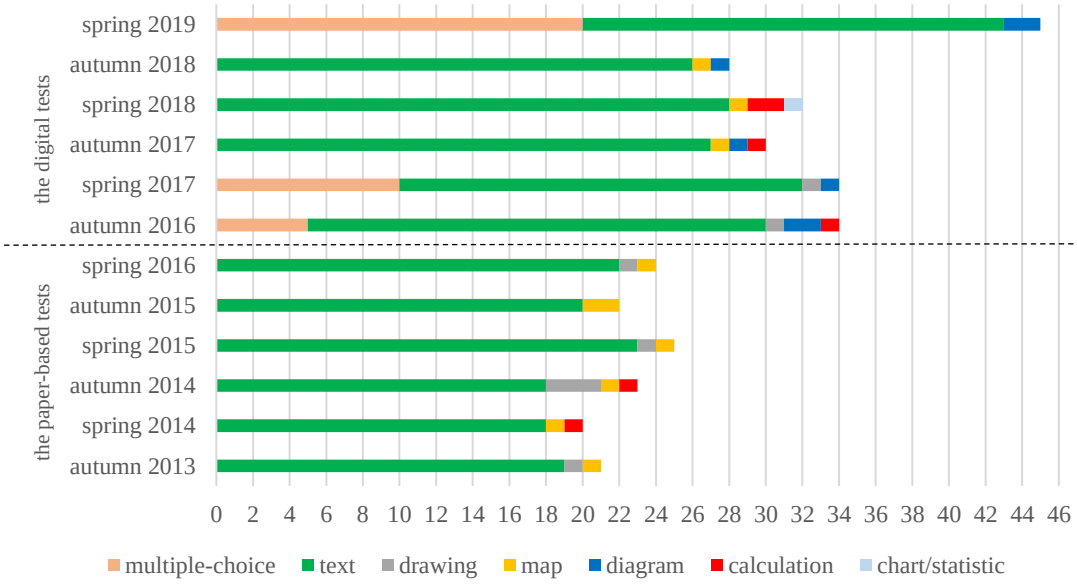

Figure 3. The number and type of geography test assignments in the Finnish Matriculation Examination between the autumn of 2013 and the spring of 2019.

5. Discussion

The majority of the geography test questions analysed are LOCS-type questions (71 % in the paper-based tests and 70 % in the digital tests). This result supports earlier research findings in different geographical education contexts (e.g., Bijsterbosch et al., 2017; Jo & Bednarz, 2009; Mishra, 2015; Wertheim & Edelson, 2013). The majority of the questions in both forms of test in our research is grouped in the category of ‘understand’, which requires explaining how something is formed, comparing differences, explaining causalities, interpreting simple diagrams and maps, giving regional examples, and inferring or predicting something. These are all essential parts of geographical thinking and learning (see Bijsterbosch et al., 2017, p. 18; see also Favier & van der Schee, 2012).

During the digitalisation process, the questions grouped in the category of ‘analyse’ slightly increased; the questions requiring remembering decreased. This means that digital tests require a student to ponder, analyse, interpret, or use given material (e.g. text, diagrams) more frequently. In this research, analysing requires information to be combined, present causalities, and form something that cannot be interpreted directly from the given material, usually from the spatial perspectives, whereas remembering includes questions in which students have to do something by relying on their memory alone: explain concepts, draw pictures and explain, or mention or name something. Remember-type questions are usually subsections of the analysed assignments, especially in the digital tests. These questions are mainly used to support more comprehensive assignments, which means that remembering knowledge is integrated with other cognitive processes in the test questions (see Anderson et al., 2014, pp. 66–69).

The questions in both forms of test mainly emphasise conceptual and factual knowledge, especially the knowledge of theories, models, and structures. In this research, these questions include geographical knowledge about climate change, urbanisation, the distribution of human population or species or industry, state development, the state of the environment, the mining industry, refugees, accessibility, and tourism. The same kinds of themes are included in the International Charter on Geographical Education (IGU-CGE, 2016, p. 10). In the digital tests, questions requiring a knowledge of terminology (for example, of geographical concepts related to geomorphology, maps, spatial data, urbanisation, natural resources, climate, and culture) increased by nine percentage points. This is probably because the subsections of the questions requiring a student to explain concepts increased during the digitalisation process.

Changes made to the structure of the examinations during the digitalisation process are seen in our statistical analysis. Accordingly, the questions in Parts I and -II are mainly LOCS-

type; the questions in Part III are mainly HOCS-type. This show that the digital tests' structure is in line with the policy of the Geography Subject Section of the FMEB (2018). Our results show that the attached material has increased and diversified, and new types of assignment have emerged, requiring students to produce different kinds of answer during the digitalisation process. However, statistical analysis revealed that the new types of attached material and new types of assignment were related to assignments which mainly required LOC-skills and factual knowledge. This means that all the possibilities brought by digitalisation (see e.g. Collins, 2018; Favier & van der Schee, 2014; Palladino, & Goodchild, 1993; van der Schee et al., 2010) have yet to be applied. However, along with the new attached materials and types of assignment, the digital tests require more diverse and comprehensive use of material, as well as more versatile production of answers. Students are also expected to use media literacy and critical understanding of these materials, as well as the ability to refer to the used materials correctly. This indicates that the digital geography tests have the potential to give students more opportunities to demonstrate their own knowledge, skills, and capabilities in geography, like the 21st century skills (see Béneker et al., 2015; Pauw, 2015, p. 308; see also Favier & van der Schee, 2012; van der Schee et al., 2010).

In studying the examinations by combining the two dimensions of the Taxonomy Table, we conclude that Finnish geography examinations demonstrate geographical thinking (see Uhlenwinkel et al., 2017, p. 10) and to some extent also powerful geographical knowledge (see Virranmäki et al., 2019, p. 11). The questions in both forms of the test mainly emphasise analysing a knowledge of theories, models, and structures, as well as remembering a knowledge of specific details and elements, and remembering a knowledge of terminology. This means that Finnish M.E. geography test questions require not only the differentiating of relevant knowledge, and organising it to form a coherent conclusion in which causalities and

a holistic approach are required; it also requires that this is done with the help of the concepts that are a crucial part of the discipline.

However, in both forms of test, there is a lack of questions that are grouped in the category of 'apply'. This is interesting, because geography is usually understood in Finland to be a school subject which frequently requires the application of information (see e.g. FNBE, 2016). However, in our research, the category of 'apply' is usually limited to using procedures (e.g. calculations, spatial data analyses, and models or theories) in different contexts. The categories of 'understand' and 'analyse' are therefore more noticeable, and the category of 'apply' is usually associated with procedural knowledge (Anderson et al., 2014, p. 77).

There is a lack of questions in the categories of 'evaluate' and 'create', indicating that the Finnish M.E. geography tests lack questions that require critical, creative, and holistic thinking, which are all major aspects of geography education (see e.g. Béneker et al., 2015; Pauw, 2015; Uhlenwinkel et al., 2017). Likewise, Bijsterbosch et al. (2017) conclude in their research on geography questions that 'the more complex cognitive processes such as evaluating and creating are hardly assessed at all' (p. 26) (see also Jo & Bednarz, 2009; Mishra, 2015; Wertheim, & Edelson, 2013). Our results show that the categories of 'create' and 'evaluate' contain questions that are very limited in their cognitive demands, and there could be a wider variety of questions within these categories. In our research, the questions in these categories require the student to ponder the pros and cons of a phenomenon. Alternatively, they must make judgements based on an aspect related to the particular phenomenon, and the questions that requires the generation of possible outcomes or the pondering of the future direction of the geographical phenomenon. They may also have to plan or produce something new that goes beyond the students' existing knowledge. Presumably, the lack of questions requiring evaluating and creating is also a result of the

challenges and workload of assessing questions in the above-mentioned levels (see also Tsaparlis & Zoller, 2003, p.56; Zheng, Lawhorn, Lumley, & Freeman, 2008).

The lack of procedural knowledge in the tests means that the use and application of geographical skills, algorithms, techniques, and methods are not assessed as they probably should be if we are to educate young people to meet the demands of the 21st century (see e.g. Béneker et al., 2015; Pauw, 2015). The absence of metacognitive knowledge is probably due to the challenges of producing these questions for large-scale summative assessments. However, there might be some kind of metacognitive level in M.E., because the students must use self-knowledge and select the assignments to be answered and assessed. Not all assignments are obligatory.

6. Conclusions

Based on our analysis of the possible change in cognitive processes and the geographical knowledge requirements of the Finnish M.E. geography test questions between the autumn of 2013 and the spring of 2019, we have addressed four future development aspects to be considered both nationally and internationally: developing policy and practice in geography assessment and the (scientific) use of revised Bloom's taxonomy in geography education.

1. If we wish to educate young people to meet the demands of the 21st century (see Béneker et al., 2015; Pauw, 2015) and to act as global citizens (see Lambert et al., 2015; Uhlenwinkel et al., 2017), or to be able to use geographical skills and knowledge, we should place more emphasis on HOC-skills, where the literature suggest the focus of the assessment should be (e.g. Airasian & Miranda, 2002; James & Gipps 1998; Tsaparlis & Zoller, 2003; see also Kumpas-Lenk et al., 2018; Radmehr & Drake, 2018). Furthermore, the knowledge dimension should be directed more towards procedural knowledge and if possible, towards metacognitive knowledge. In Finland, the practice of planning and producing digital tests presents significant opportunities to do this, because using digital technologies can guide

students' learning towards HOC-skills (see e.g. Collins, 2018; Favier, & van der Schee, 2014; Palladino, & Goodchild, 1993).

2. However, we suggest that there is a need to discuss how much emphasis should be placed on HOC-skills. Since LOCS are an important basis for employing HOCS, we suggest a discussion and consideration of their ideal distribution. During the study's examination period, the division was approximately 30 per cent in HOCS and 70 per cent in LOCS. Overall, we need international dialogue between geography departments, researchers, and teacher-educators in universities on the one hand, and geography school-teachers on the other, about the preferred policy and practice of geography assessment. We need to critically evaluate the kind of cognitive processes and geographical knowledge we want to see assessed in the geography tests, because assessment affects teaching and learning (Baird et al., 2017, p. 340; see also Bijsterbosch et al., 2017; Chang, & Seow, 2018; Ormond, 2019). As Uhlenwinkel et al. (2017) point out, 'an international debate necessarily asks for an open discourse in which while trying to reach the same goals, each country can learn from the other' (p. 7). In Finland, the importance of the Finnish M.E. will increase in 2020, when the majority of students will be accepted into higher education based on their success in it. It is therefore important to consider the universities' perspective concerning the kind of geographical knowledge and cognitive processes the tests should require.

3. We discovered that the Finnish digital tests required more analysis as well as comprehensive use and production of materials, and that the Part III -assignments required more HOCS-skills than other parts of the tests. Therefore, geography teachers and students should be aware of these aspects. We need training for both in-service and pre-service teachers on the principles of assessment and changes in policy concerning the requirements (see also DeLuca & Johnson, 2017; Wertheim & Edelson, 2013, p. 15; Yasar, 2009, p. 64). As Ormond (2019, p. 18) points out, assessment is not always transparent, e.g. teachers and

students may not know what assessment requires. However, we should note that assessment can improve teaching and learning (Baird et al., 2017; Stoltman et al., 2014, p. 193; Wertheim & Edelson, 2013, p. 15). As Wertheim and Edelson (2013) point out, 'well-designed assessments can be used to monitor students' progress, inform instructional decisions, and pinpoint effective programs and practices' (p. 15). Therefore, teachers especially should know the kinds of cognitive process and knowledge dimension assessment questions require to ensure they can provide high-quality teaching (see Anderson, 2005, p. 111). Although the geography tests do not emphasise HOC-skills, this does not mean that teachers do not emphasise these skills in their teaching (see also Bijsterbosch et al., 2017, p. 27). However, we should be careful to avoid 'teaching to test' (see e.g. Bijsterbosch et al., 2017; Torrance, 2011). We should also be careful to avoid using the procedures and processes of assessment 'as a way to frame the curriculum and drive the reform of schooling' (Torrance, 2011, p. 459).

4. We acknowledge that geographical education settings and its aims vary across the globe (see e.g. Butt & Lambert, 2014; Brooks et al., 2018; Uhlenwinkel et al., 2017), and our results should be interpreted in the Finnish education context. Caution is urged when considering the results, yet we conclude that our results are internationally noteworthy. We have created a framework for analysing the change in geography test questions' cognitive and knowledge requirements during the digitalisation process by applying revised Bloom's taxonomy in the context of Finnish M.E. Our research contributes new knowledge to the suggested need for more research into geographical cognitive skills and knowledge (see Lane, & Bourke, 2017; Leat et al., 2012; Wertheim, & Edelson, 2013). We have explained (Tables 2 and 3) how we have used the Taxonomy Table in the context of geography, and we urge other scientists to use it in their national educational contexts. By using the same kind of framework, the known similarities and differences between geography education in different countries will be acknowledged.

We need more information and dialogue internationally on how to utilise revised Bloom's taxonomy in research, and teaching and assessing subjects like geography. Although Krathwohl (2002) states that 'it gives commonly understood meaning to objectives classified in one of its categories' (p. 218), it is also a subjective statement as to how to interpret the Taxonomy Table (see Anderson et al., 2014, p. 33). Furthermore, there is a need for more research to examine international differences between different national large-scale assessments to produce knowledge of assessed cognitive processes and knowledge demands, and the preferred direction of assessments and education to increase young people's capability as future global citizens.

References

Airasian, P. W., & Miranda H. (2002). The Role of Assessment in the Revised Taxonomy. *Theory Into Practice, 41*(4), 249–254. https://doi.org/10.1207/s15430421tip4104_8

Anderson, L. W. (2005). Objectives, Evaluation, and the Improvement of Education. *Studies in Educational Evaluation, 31*(2–3), 102–113. https://doi.org/10.1016/j.stueduc.2005.05.004

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. (2014). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's.* Edinburgh, England: Pearson Education Limited.

Baird, J.-A., Andrich, D., Hopfenbeck, T. N., & Stobart, G. (2017). Assessment and learning: fields apart? *Assessment in Education: Principles, Policy & Practice*, *24*(3), 317–350. https://doi.org/10.1080/0969594X.2017.1319337

Béneker, T., Palings, H., & Krause, U. (2015). Teachers envisioning future geography education at their schools. *International Research in Geographical and Environmental Education, 24*(4), 355–370. http://doi.org/10.1080/10382046.2015.1086102

Bijsterbosch, E., van der Schee, J., & Kuiper, W. (2017). Meaningful learning and summative assessment in geography education: an analysis in secondary education in the Netherlands. *International Research in Geographical and Environmental Education*, *26*(1), 17–35. https://doi.org/10.1080/10382046.2016.1217076

Brooks, C., Qian, G., & Salinas-Silva, V. (2018). Geographical Union – Commission for Geography Education What next for Geography Education ? *A perspective from the International Geographical Union – Commission for Geography Education, (May 2017)*, 5–15. http://doi.org/10.4458/8579-01

Butt, G., & Lambert, D. (2014). International perspectives on the future of geography education: an analysis of national curricula and standards. *International Research in Geographical and Environmental Education, 23*(1), 1–12. http://doi.org/10.1080/10382046.2013.858402

Chang, C. H., & Kidman, G. (2019). Curriculum, pedagogy and assessment in geographical–for whom and for what purpose? *International Research in Geographical and Environmental Education*, 28(1), 1–4. https://doi.org/10.1080/10382046.2019.1578526

Chang, C. H., & Seow, T. (2018). Geographical Education That Matters – A Critical Discussion of Consequential Validity in Assessment of School Geography. *Geographical Education*, 31, 31–40.

Collins, L. (2018). The Impact of Paper Versus Digital Map Technology on Students' Spatial Thinking Skill Acquisition. *Journal of Geography, 117*(4), 137–152. http://doi.org/10.1080/00221341.2017.1374990

Commission of Geography Education of the International Geographical Union [IGU-CGE] (2016). International Charter on Geographical Education. Retrieved from *http://www.igu-cge.org/wp-content/uploads/2019/03/IGU_2016_eng_ver25Feb2019.pdf*

DeLuca, C., & Johnson, S. (2017). Editorial: Developing assessment capable teachers in this age of accountability. *Assessment in Education: Principles, Policy & Practice*, *24*(2), 121–126. https://doi.org/10.1080/0969594X.2017.1297010

Favier, T. T., & van der Schee, J. (2012). Exploring the characteristics of an optimal design for inquiry-based geography education with Geographic Information Systems. *Computers & Education, 58*(1), 666–677. https://doi.org/10.1016/j.compedu.2011.09.007

Favier, T. T., & van der Schee, J. (2014). The effects of geography lessons with geospatial technologies on the development of high school students' relational thinking. *Computers & Education, 76*, 225–236. http://doi.org/10.1016/j.compedu.2014.04.004

Finnish Matriculation Examination Board. (2017a February 9). Reaaliaineiden paperisten kokeiden määräykset. [Regulations of the paper Matriculation Examination of humanities and natural sciences]. Retrieved from *https://www.ylioppilastutkinto.fi/images/sivuston_tiedostot/Ohjeet/Koekohtaiset/maaraykset_reaaliaineet_090217.pdf*

Finnish Matriculation Examination Board. (2017b March 10). Reaaliaineiden digitaalisten kokeiden määräykset. [Regulations of the digital Matriculation Examination of humanities and natural sciences]. Retrieved from *https://www.ylioppilastutkinto.fi/images/tiedostot/reaaliaineiden_digitaalisten_kokeiden_maaraykset_2017.pdf*

Finnish Matriculation Examination Board. (n.d.). Sähköinen ylioppilastutkinto reaaliaineet. [Digital Matriculation Examination in the humanities and natural sciences]. Retrieved from *https://www.ylioppilastutkinto.fi/images/sivuston_tiedostot/Sahkoinen_tutkinto/fi_sahkoinen_reaali.pdf*

Finnish National Board of Education. (2016). *National Core Curriculum for General Upper Secondary Schools* 2015. Helsinki: Next Print Oy.

Geography Subject Section of the Finnish Matriculation Examination Board. (2018 January 10). Tiedote maantieteen opettajille ja opiskelijoille [Notice to geography teachers and students]. Retrieved from

https://www.ylioppilastutkinto.fi/images/sivuston_tiedostot/Sahkoinen_tutkinto/ge_tiedote_fi.pdf

Hooghuis, F., van der Schee, J., van der Velde, M., Imants J., & Volman, M. (2014). The adoption of Thinking Through Geography strategies and their impact on teaching geographical reasoning in Dutch secondary schools. *International Research in Geographical and Environmental Education, 23*(3), 242–258. https://doi.org/10.1080/10382046.2014.927168

James, M., & Gipps, C. (1998). Broadening the basis of assessment to prevent the narrowing of learning. *The Curriculum Journal, 9*(3), 285–297. https://doi.org/10.1080/0958517970090303

Jo, I., & Bednarz, S. W. (2009). Evaluating geography textbook questions from a spatial perspective: Using concepts of space, tools of representation, and cognitive processes to evaluate spatiality. *Journal of Geography*, *108*(1), 4–13. https://doi.org/10.1080/00221340902758401

Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An overview. *Theory into Practice, 41*(4), 37–41. https://doi.org/10.1207/s15430421tip4104_2

Kumpas-Lenk, K., Eisenschmidt, E., & Veispak, A. (2018). Does the design of learning outcomes matter from students' perspective? *Studies in Educational Evaluation, 59*, 179–186. https://doi.org/10.1016/j.stueduc.2018.07.008

Lambert, D. (2017). Powerful disciplinary knowledge and curriculum futures. In N. Pyyry, L. Tainio, K. Juuti, R. Vasquez, & M. Paananen (Eds.) (2017) *Changing subjects, changing pedagogies: diversities in school and education*. (pp. 14–31). Publications of the Finnish Research Association for Subject Didactics: Studies in Subject Didactics 13

Lambert, D., Solem, M., & Tani, S. (2015). Achieving Human Potential Through Geography Education: A Capabilities Approach to Curriculum Making in Schools. *Annals of the Association of American Geographers*, *105*(4), 723–735. https://doi.org/10.1080/00045608.2015.1022128

Lane, R., & Bourke, T. (2017). Assessment in geography education: a systematic review. *International Research in Geographical and Environmental Education*, *28*(1), 22–36. https://doi.org/10.1080/10382046.2017.1385348

Leat, D., Thomas, U., & Reid, A. (2012). The Epistemological Fog in Realising Learning to Learn in European Curriculum Policies. *European Educational Research Journal, 11*(3), 400–412. https://doi.org/10.2304/eerj.2012.11.3.400

Liu, Y., Bui, E. N., Chang, C.-H., & Lossman, H. G. (2010). PBL-GIS in Secondary Geography Education: Does It Result in Higher-Order Learning Outcomes? *Journal of Geography, 109*(4), 150–158. http://doi.org/10.1080/00221341.2010.497541

Maude, A. (2015). What might powerful geographical knowledge look like*? Geography, 101*(2), 70–76.

Maude, A. (2016). What is powerful knowledge and can it be found in the Australian geography curriculum? *Geographical Education, 28*, 18 –26.

Maude, A. (2018). Geography and powerful knowledge: A contribution to the debate. *International Research in Geographical and Environmental Education*, *27*(2), 179–190. https://doi.org/10.1080/10382046.2017.1320899

Mishra, R. K. (2015). Mapping the knowledge topography: A critical appraisal of geography textbook questions. *International Research in Geographical and Environmental Education, 24*(2), 118–130. https://doi.org/10.1080/10382046.2014.993170

Nordmark, M. (2014). Digitalt skrivande i qymnasieskoland svenskundervisning. En ämnesdidiaktisk studie av skrivprocessen. [Digital writing in the teaching of Swedish at upper secondary school: A study of the writing process in the field of subject didactics.] (Doctoral Dissertation). Retrieved from *https://www.diva-portal.org/smash/get/diva2:1151399/FULLTEXT01.pdf*

Ormond, B. M. (2019). The Impact of standards-based assessments on knowledge for history education in New Zealand. *Assessment in Education: Principles, Policy & Practice, 26(*2), 143–165. https://doi.org/10.1080/0969594X.2018.1432564

Palladino, S., & Goodchild, M. (1993). A place for GIS in the secondary schools? Lessons from the NCGIA secondary education project. *Geo Info Systems, 3*(4) 45–49.

Pauw, I. (2015). Educating for the future: the position of school geography. *International Research in Geographical and Environmental Education*, *24*(4), 307–324. https://doi.org/10.1080/10382046.2015.1086103

Pellegrino, J. W. (2017). The two disciplines problem – 'it's like Déjà vu all over again!'. *Assessment in Education: Principles, Policy & Practice*, *24*(3), 359–368. https://doi.org/10.1080/0969594X.2017.1326876

Radmehr, F., & Drake, M. (2018). An assessment-based model for exploring the solving of mathematical problems: Utilizing revised Bloom's taxonomy and facets of metacognition. *Studies in Educational Evaluation, 59*, 41–51. https://doi.org/10.1016/j.stueduc.2018.02.004

Rautapuro, J., & Harjunen, E. (2017) Mode effect in large-scale assessment. In N. Pyyry, L. Tainio, K. Juuti, R. Vasquez, & M. Paananen (Eds.) (2017) *Changing subjects, changing*

pedagogies: diversities in school and education. (pp. 260–274). Publications of the Finnish Research Association for Subject Didactics: Studies in Subject Didactics 13.

Slater, F., Graves, N., & Lambert, D. (2016). Editorial. *International Research in Geographical and Environmental Education, 25*(3), 189–194. https://doi.org/10.1080/10382046.2016.1155321

Stoltman, J., Lidstone, J., & Kidman, G. (2014). Geography and international assessment: opportunity or distraction. *International Research in Geographical and Environmental Education, 23*(3), 193–196. https://doi.org/10.1080/10382046.2014.927172

Tikkanen, G., & Aksela, M. (2012). Analysis of Finnish chemistry Matriculation Examination questions according to Cognitive Complexity. *Nordina, 8*(3), 258–268. https://doi.org/10.5617/nordina.532

Torrance, H. (2011). Using Assessment to Drive the Reform of Schooling: Time to Stop Pursuing the Chimera? *British Journal of Educational Studies, 59*(4), 459–485. https://doi.org/10.1080/00071005.2011.620944

Torrance, H. (2017). Blaming the victim: assessment, examinations, and the responsibilisation of students and teachers in neo-liberal governance. *Discourse: Studies in the Cultural Politics of Education, 38*(1), 83–96. https://doi.org/10.1080/01596306.2015.1104854

Tsaparlis, G., & Zoller, U. (2003). Evaluation to higher vs. lower-order cognitive skill-type examinations in chemistry: implications for university in-class assessment and Examinations. *University Chemistry Education, 7*(2), 50–57. https://doi.org/10.18844/cjes.v1i1.67

Uhlenwinkel, A., Béneker, T., Bladh, G., Tani, S., & Lambert, D. (2017). GeoCapabilities1 and curriculum leadership: balancing the priorities of aim-based and knowledge-led curriculum thinking in schools. *International Research in Geographical and Environmental Education, 26*(4), 327–341. https://doi.org/10.1080/10382046.2016.1262603

Valtioneuvoston kanslia (2011 June 22). Pääministeri Jyrki Kataisen hallituksen hallitusohjelma [Prime Minister's Office Finland, Government Program of Prime Minister Jyrki Katainen's Government]. Retrieved from *http://valtioneuvosto.fi/documents/10184/147449/Kataisen+hallituksen+ohjelma/81f1c20f-e353-47a8-8b8f-52ead83e5f1a/Kataisen+hallituksen+ohjelma.pdf*

Van Der Schee, J., Notté, H., & Zwartjes, L. (2010). Some thoughts about a new international geography test. *International Research in Geographical and Environmental Education, 19*(4), 277–282. https://doi.org/10.1080/10382046.2010.519148

Virranmäki, E., Valta-Hulkkonen, K., & Rusanen, J. (2019). Powerful Knowledge and the significance of teaching geography for in-service Upper Secondary geography teachers.

International Research in Geographical and Environmental Education, 28(2), 103–117. https://doi.org/10.1080/10382046.2018.1561637

Wertheim, J. A., & Edelson, D. C. (2013). A Road Map for Improving Geography Assessment. *Geography Teacher*, *10*(1), 15–21. https://doi.org/10.1080/19338341.2012.758044

Yang, D. (2013). Comparing Assessments Within Junior Geography Textbooks Used in Mainland China. *Journal of Geography, 112*(2), 58–67. https://doi.org/10.1080/00221341.2011.648211

Yasar, O. (2009). A Comparative Analysis of Assessment and Evaluation Exercises Included in Geography Textbooks Written According to the 2005 Secondary Education Geography Curriculum and Textbooks of the Former Curriculum in Turkey. *International Journal of Progressive Education, 5*(1), 45–68.

YLE. (2019 August 19). Yo-kokeet maantiede. [The Matriculation Examination: Geography]. Retrieved from https://yle.fi/aihe/artikkeli/2015/12/15/yo-kokeet-maantiede

Zheng, A. Y., Lawhorn, J. K., Lumley, T., & Freeman, S. (2008). Application of Bloom's taxonomy debunks the "MCAT myth". *Science 319*, 414–415. https://doi.org/10.1126/science.1147852

Zoller, U., & Pushkin, D. (2007). Matching Higher-Order Cognitive Skills (HOCS) promoting goals with problem-based laboratory practice in a freshman organic chemistry course. *Chemistry Education Research and Practice, 8*(2), 153–171. https://doi.org/10.1039/B6RP90028C