Peer-reviewed academic journal

Innovative Issues and Approaches in Social Sciences

IIASS – VOL. 14, NO. 1, JANUARY 2021

Innovative Issues and Approaches in Social Sciences

IIASS is a double blind peer review academic journal published 3 times yearly (January, May, September) covering different social sciences: political science, sociology, economy, public administration, law, management, communication science, psychology and education.

IIASS has started as a SIdip – Slovenian Association for Innovative Political Science journal and is being published by ERUDIO Center for Higher Education.

Typeset
This journal was typeset in 11 pt. Arial, Italic, Bold, and Bold Italic; the headlines were typeset in 14 pt. Arial, Bold

Abstracting and Indexing services
COBISS, International Political Science Abstracts, CSA Worldwide Political Science Abstracts, CSA Sociological Abstracts, PAIS International, DOAJ, Google scholar.

Publication Data:
ERUDIO Education Center

Innovative issues and approaches in social sciences,
2021, vol. 14, no. 1

ISSN 1855-0541

Additional information: www.iiass.com

THE TAXONOMY OF THE QUESTIONS IN THE ITALIAN NATIONAL ASSESSMENT OF KNOWLEDGE OF MATHEMATICS

Daniel Doz[5]

Abstract

In Italy, all grade 10 students are required to take the national assessment of knowledge of mathematics, which is prepared by the INVALSI Institute. No official data about any taxonomic level of the questions in these assessments has been published on the Institute's website. In the present work, we analyzed seven INVALSI examinations for school with Slovene as language of instruction and we focused on the Gagne's taxonomic level of each individual question in the assessments. The most frequent category in national assessments is "Routine procedural knowledge", followed by "Basic and conceptual knowledge". We found that, even though the interest in problem-solving activities has increased in the past years, the taxonomic level "Problem-solving knowledge" is the less frequent. Moreover, we wanted to analyze the distribution of the different taxonomic levels among the question typologies (open- and closed-type questions) and we found that questions from lower taxonomic levels are more likely to be closed-type, while "Problem-solving questions" are more likely to be open-type. Furthermore, we were interested in analyzing the distribution of taxonomic levels among the four topic dimensions "Geometry", "Data and prevision", "Numbers and quantities" and "Relations and functions". We found that the taxonomic level "Problem-solving knowledge" are more present in the categories "Number and quantities" and "Relations and functions".

Key words: mathematics, national examinations, taxonomy, Gagne.

DOI: http://dx.doi.org/10.12959/issn.1855-0541.IIASS-2021-no1-art4

---

[5] Faculty of Education, Universtiy of Primorska, Slovenia, doz_daniel@yahoo.it

Introduction

There has been a change of interest in mathematics education from a mere teaching of mechanical procedures to a focus on conceptual knowledge and strategies of solving problems, which are transferable also to other real-life situations (Cotič, 2010; Vos, 2018; Altay et al., 2017). This shift of interest of how students learn mathematics is also a consequence of the shift in the theory of learning from behaviorism to constructivism (Rosli et al., 2013).

Although mathematical competence is achieved by developing both conceptual and procedural knowledge (Rittle-Johnson & Schneider, 2015; Zuya, 2017), procedural knowledge has decreased in importance, whereas the necessity of conceptual knowledge has increased. Modern strategies in teaching mathematics have brought some changes also in the evaluation of this subject (Cotič, 2010). Hence, it is also important that mathematics teachers and theoreticians of mathematics educations are able to critically analyze, understand and use the different kinds of acquiring new information, also in order to efficiently evaluate student's knowledge. Assessing students' understanding of mathematics in a traditional way seems to be particularly difficult (Anderson & Puckett, 2003), since traditional tests do only focus on students' "mechanical" mathematical skills and evaluate merely the procedures (Kulm, 1994). Rosli and colleagues (2013) affirmed that, from a point of view of learning and teaching mathematics, the method of problem-solving has become a central approach in the classroom, since it permits the teacher to understand how students learn mathematics. The authors have, however, emphasized that the methods to evaluate students' learning and understanding of mathematics, that are based on a meaningful assessment of constructivist problems, are still underdeveloped.

Cotič (2010) stated that the constant assessment of the students' understanding plays a core role in the learning and teaching process. The evaluation of students' knowledge can help them to think critically about the studied topics and eventually to clarify their doubts and misunderstandings. On the other hand, assessments provide vital feedback information both to students and teachers (Felda, 2018). Students can, indeed, understand which their strong and weak points are; moreover, they can comprehend what did they understand (Guskey, 2019). Teachers can get a more objective view about their teaching style and their classes, both from a point of view of contents and taxonomy.
There are different factors that contribute to an assessment's quality, such as the typology of the assessment itself (Cotič, 2010). Traditional assessments, where students must solve an equation, draw a geometrical shape, or compute an expression, are hence an insufficient

way of evaluating students' knowledge and competences. On the contrary, the need of evaluating also students' problem-solving abilities and connecting the theoretical knowledge to a concrete (real-world) situation, is increasing (Lester Jr., 2013; Sumarna & Herman, 2017; Yuanita et al., 2018). It is hence important to assess different levels of students' understanding of mathematics. Students' educational learning objectives can be classified into levels of complexity using a taxonomic scale, which indicates the complexity level of a certain problem or task (Thompson, 2008).

For what it concerns the taxonomy levels of national assessments of mathematical knowledge in Italy, promoted and organized by the INVALSI Institute, no information about the taxonomy levels of the questions is known. In contrast with the situation with the national assessments in Slovenia, where the percentage of the questions from each taxonomic level, mathematical topic and question type are specified in some national documents (see, e.g. RIC, 2005, 2015, 2019), the INVALSI institute does not specify the quantity of problem from each taxonomic level.

In this work we aim to investigate which level of the Gagne's classification of learning is the most present in the INVALSI assessments of knowledge. In particular, in this paper we would like to present which category (from the Gagne's taxonomy) is the most common in the INVALSI examinations of mathematics for grade 10 students of Italian high schools with Slovene as language of instruction. We are, furthermore, interested whether specific question typologies (open- or closed-type) are more present in some taxonomic level, and which taxonomic level is more present among the questions about a specific mathematical topic.

Theoretical Framework

Why are the National Assessments of Mathematics useful?

In order to constantly monitor students' achievements, and hence the improvements in the quality of the schooling system, many countries use the national assessments of knowledge (Cankar, 2008; Bansilal, 2017; Kartianom & Mardapi, 2017; Felda, 2018). These assessments do only evaluate students' cognitive achievements and only in some specific subjects (Sulistyaningsih & Sugiman, 2016), but they have the big advantage to be an objective and reliable kind of measuring such achievements (Cankar, 2008). National assessments of knowledge are hence an objective instrument that can be used to evaluate what students know, which are the most common mistakes, and which are the most frequent doubts that students have. Furthermore, the results of national assessments of knowledge can help educators and teachers of a specific subject to understand, how is it possible to improve (or correct) their teaching methods (Cakar, 2008; Felda, 2018).

Students' final grades in mathematics do not necessarily correspond to the marks that the same students would get in the national assessment of knowledge. Some students that are somehow "bad" at school, would result very good in a national assessment of knowledge; vice versa, a student with excellent grades in mathematics would barely "pass" the national assessment of knowledge (Felda, 2018).

The Italian context

The national assessments of knowledge in Italy are organized by the National System for the Evaluation (SNV), which works inside the National Institute for the Evaluation of the Educative System for Instruction and Formation (INVALSI). The INVALSI institute has the role to investigate and assess students' knowledge (directive MIUR 76/2009) and the quality of the Italian schooling system (D. Lgs. n. 286/2004). As it is stated in the Quadro di Riferimento (2017, 2018), the INVALSI institute assess students' knowledge of Italian (or minority) language, mathematics and English among grade 2, 5, 8 and 10 students, while from the school year 2018/19 it has become mandatory to evaluate the knowledge of the Italian language, mathematics and English also among grade 13 students.

In the Quadro di Riferimento (2017, 2018) and in the work of Martignone (2016) it is stated that the INVALSI institute prepares the national assessments of knowledge for the primary school considering the national curriculum indications present in the document "Indicazioni Nazionali per il curricolo della scuola dell'infanzia e del primo ciclo di istruzione", while the national examinations for the middle and high school are based on the national curricula indications "Indicazioni Nazionali e Linee guida per le scuole secondarie di secondo grado". The topics that the INVALSI Institute considers in order to prepare the national assessments of knowledge are hence known, since the questions that are asked in each examination should be referred to the specific national curricula indication. No data about how many questions about each topic the students would get in their national assessments is known in advance. In the Quadro di Riferimento (2017, 2018) it is only stated that the questions are equally distributed among the various mathematical topics.

Also, no data about the taxonomic level of the questions is known. In the Quadro di Riferimento (2017) it is stated that the proposed problems have the aim to assess two dimensions, similarly to the TIMSS (2019) assessments:

(1) the cognitive dimension: in the TIMSS (2019) document, the cognitive domains are defined as the processes that we expect the students would apply while working on a mathematical problem. TIMSS distinguish among three cognitive domains: knowing, applying and reasoning. In the knowledge domain we consider the facts, concepts and procedures that students should know. The applying domain regards the capacity of students to apply their knowledge and learnt concepts to solve a certain problem or to answer a specific question. The reasoning domain includes all those strategies that students should know in order to analyze unfamiliar or complex situations. In Table 1 (see TIMSS, 2019, p. 14) we might see which cognitive domains are present in assessments for grade 4 and 8. From Table 1 we might understand that in grade 8 there are more reasoning-related problems than in grade 4.

Table 1: The cognitive domains in the TIMSS (2019) assessments.

Cognitive domains	Percentages	
	Grade 4	Grade 8
Knowing	40%	35%
Applying	40%	40%
Reasoning	20%	25%

(2) the topic dimension: it represents the dimension of the four mathematical topics that are assessed in the examinations (numbers, algebra, geometry and data & probability). In Table 2 we present the percentages of each topic domain present in the TIMSS (2019) examinations.

Table 2: The topic domains in the TIMSS (2019) assessments for grade 8.

Content domains	Percentages
Number	30%
Algebra	30%
Geometry	20%
Data and Probability	20%

In the Quadro di Riferimento (2017) it is stated that it is important to consider two key aspects about mathematics itself:

(1) an applicative part, where students need to model natural and everyday phenomena through mathematical models, apply the knowledge of mathematics to solve everyday problems and interpret the reality;
(2) a more theoretical part, based on an internal reflection and growth, a critical analysis of the reality and the mathematics itself.

Students should, consequently, understand the concepts and elementary methods of mathematics that are relevant both for the development of the mathematical knowledge itself and for a description of simple natural and physical phenomena. The teaching of mathematics should help the students to acquire the knowledge and the competences that could help them to critically analyze the modern world.

In the Decree 211/2010, published in the Gazzetta Ufficiale (2010), it is stated that mathematics teachers should select both theoretical and applicable exercises. In particular, the problems that students would solve, should be inherent to some theoretical topics, but also to some more practical issues, especially related to economics, social sciences, technology, physics, natural sciences and to the real world. Furthermore, students should be able to use a formal language and to prove some basic theorems, they should have acquired also the abilities and competences to solve mathematical, statistical and probability problems, in order to understand better the other scientific disciplines. In the same Decree it is also underlined the importance that real-world applications and experiences have in the teaching of mathematics: teachers should present this subject as a product of a long cultural development and, at the same time, as an applicable discipline.

In the Quadro di Riferimento (2017, 2018) it is stated that standardized assessments cannot substitute the teacher in the process of evaluation of the students' knowledge. The feedback that the teacher would get, should represent only an additional information about the level of knowledge and competences that the students have achieved, since the national assessments of knowledge in Italy:

(i) cannot measure nor evaluate the achievement of some metacognitive or non-cognitive goals, such as the development of a positive attitude towards the subject, or the students' proficiency in using technology to implement the theoretical lessons;
(ii) cannot evaluate students' abilities in producing conjectures, argument and execute complex proofs of theorems. Moreover, the INVALSI examinations cannot assess the students' abilities to solve some complex problems that

require several steps to be solved. The INVALSI examinations also cannot evaluate the students' abilities to model more complex real-world situations;

(iii) are objective and they do not depend on the person that corrects them, but they do not take anyhow into account the affective and social component that is present in a teacher's evaluation. Teachers know their students' social, economic and cultural background, which might help them to evaluate students in a more complete way, since they can consider also students' progresses.

In order to evaluate the students' competences and abilities in mathematics, the INVALSI Institute has classified the questions into three different categories: solving problems, reasoning and knowing (Quadro di Riferimento, 2017; INVALSI, 2017). The dimensions of solving problems and reasoning are connected (INVALSI, 2017), since the process of argumentation is a typical problem-solving process; on the other hand, while working on problem-solving, an intermediate validation argumentative process is also required. Both dimensions require that the students know some concepts, signs, algorithms and solving techniques. Moreover, students should show their work and thinking, check the validity of their results and adequately justify their answer. The tasks that require some argumentation skills from the students are present, in the INVALSI assessments, as three possible problems:

(1) the student should choose the correct answer and individuate the correct argumentation for the validity of such answer;
(2) the student should choose the correct answer and give the correct argumentation for the validity of such answer;
(3) the student has to formulate a hypothesis and verify its validity, conjecture and prove his/her conjecture.

The authors of the report INVALSI (2017) affirm that the production of the second and third kind of questions is particularly hard.

In Table 3 we present the frequency of problems of a specific content dimension and specific cognitive dimension in the INVALSI assessment of knowledge of mathematics from the school year 2015-16 for grade 10 students. We can notice that the most frequent cognitive dimension is "Knowing" (42,9%), followed by an almost equally frequent "Solving problems" (40,8%). The category that is the less frequent is "Reasoning" (16,3%).

Table 3: The topic and cognitive domains in the INVALSI 2016 assessment of mathematics for grade 10 students (INVALSI, 2016, p. 23).

Content domains	Cognitive content			Total
	Knowing	Solving problems	Reasoning	
Numbers	9 (69,2%)	3 (23,1%)	1 (7,7%)	13 (100,0%)
Space and shapes	7 (63,6%)	3 (27,3%)	1 (9,1%)	11 (100,0%)
Data and prevision	3 (20,0%)	11 (73,3%)	1 (6,7%)	15 (100,0%)
Relations and functions	2 (20,0%)	3 (30,0%)	5 (50,0%)	10 (100,0%)
Total	21 (42,9%)	20 (40,8%)	8 (16,3%)	49 (100,0%)

In Table 4 we present the frequency of problems of a specific content dimension and specific cognitive dimension in the INVALSI assessment of knowledge of mathematics from the school year 2016-17 for grade 10 students. From the analyzed percentages we might find out that the most common cognitive content is “Knowing” (49,1%), followed by “Solving problems” (34,0%) and “Reasoning” (17,0%). From the analyzed data we might conclude that the INVALSI examination of mathematic for grade 10 students in the school year 2016-17 is mostly focused on knowing terms and concepts, rather than to argument the given answers and solving problems. From the Tables 3 and 4 we conclude that “Reasoning” is commonly the less frequent dimension that is assessed by the INVALSI assessments in the years 2016 and 2017.

Table 4: The topic and cognitive domains in the INVALSI 2017 assessment of mathematics for grade 10 students (INVALSI, 2017, p. 24).

Content domains	Cognitive content			Total
	Knowing	Solving problems	Reasoning	
Numbers	5 (27,8%)	7 (38,9%)	6 (33,3%)	18 (100,0%)
Space and shapes	7 (77,8%)	1 (11,1%)	1 (11,1%)	9 (100,0%)
Data and prevision	6 (54,5%)	5 (45,4%)	0 (0,0%)	11 (100,0%)
Relations and functions	8 (53,3%)	5 (33,3%)	2 (13,3%)	15 (100,0%)
TOTAL	26 (49,1%)	18 (34,0%)	9 (17,0%)	53 (100,0%)

In Table 1 we presented the cognitive domains in TIMSS (2019). We would like to compare these percentages with the ones we found in Tables 3 and 4. We will consider the domain “Reasoning” from the TIMSS examinations as equivalent to “Reasoning” in the INVALSI assessments and the TIMSS category “Applying” with the INVALSI “Solving problems” domain, while there are some differences between these two pairs of categories. We present the results in Table 5.

Table 5: The cognitive domains in the TIMSS (2019), INVALSI (2016) and INVALSI (2017) assessments.

Cognitive domains	Percentages			
	TIMSS (2019)		INVALSI (2016)	INVALSI (2017)
	Grade 4	Grade 8		
Knowing	40%	35%	42,9%	49,1%
Applying	40%	40%	40,8%	34,0%
Reasoning	20%	25%	16,3%	17,0%

For the sake of comparison, we can analyze the similarities and the differences between the Grade 8 TIMSS examinations and the Grade 10 INVALSI assessments.

From the analysis in Table 5 we can notice that there are some differences between the distributions of problems in the TIMSS and INVALSI examinations for what it concerns the cognitive domains. We can observe that the INVALSI assessments have generally more

"Knowing" problems that the TIMSS examinations, while the "Reasoning" problems are more present in the TIMSS assessments than in the INVALSI problems.

Gagne's Classification of Knowledge

Gagne's contributions to educational psychology are very significant (Zhang et al., 2010). In his researches (Gagne, 1972, 1977, 1984), Gagne classified the learning type in accordance to the learning results. Students' achievements, in accordance to the levels of knowledge, can be described with a taxonomic scale, which is divided in three parts (Cotič, 2010):

(1) basic and conceptual knowledge: in this dimension we include the knowledge of terms and concepts, and also the ability to recall the previously acquired knowledge. Conceptual knowledge is the understanding of facts and concepts. Among the classical elements that are included in the basic and conceptual knowledge we might find:
   a. knowledge of the individual elements: fractographical knowledge, knowing isolated information, reproductive knowledge;
   b. knowledge of specific facts: knowledge of definitions, formulas, axioms, basic relations among mathematical objects, knowledge of specific mathematical terminology and symbols (such as "parallelism", "equation", "function" etc.), knowing how to classify these in categories, knowing how to represent two objects (e.g. two congruent right triangles can form a rectangle);

(2) procedural knowledge: the dimension of procedural knowledge does include the identification and the correct knowledge of procedures and algorithms. We can divide it in:
   a. routine procedural knowledge;
   b. complex procedural knowledge.

   Procedural knowledge is based mainly on the execution of routine steps, the application of rules and formulas, solving elementary problems with limited quantity of information, the recognition and correct application of algorithms and procedures, the usage of rules, the correct choice of solving algorithms and procedures, solving complex problems (exercises composed by different tasks with multiple initial data);

(3) problem-solving knowledge: this domain is mainly focused on the application of the learnt mathematical methods in new situations, combining multiple rules and concepts. The abilities of using conceptual and procedural knowledge is hence essential in this dimension. This kind of knowledge require that the students

recognizes the problem, its formulation and its questions, that knows how to analyze the initial data (if there are not enough data to solve the problem, or whether the problem has too many initial data, etc.), must find a correct strategy of solving the problem (with the help of communication, operative and cognitive processes), apply their knowledge to a different context. At this level, metacognitive processes are also a key element, since they help students to judge if they are applying the correct formulas in a specific context.

In this research we will consider four Gagne's taxonomic levels:
(1) basic and conceptual knowledge;
(2) routine procedural knowledge;
(3) complex procedural knowledge;
(4) problem-solving knowledge.

Empirical Research

Aims of the Research

In Tables 3 and 4 we presented the frequency and the percentages of each cognitive domain present in the INVALSI assessments of knowledge of mathematics for grade 10 students. We compared these results with that of the TIMSS examinations and we found that the distribution of the problems among the three main categories "Knowing", "Applying" and "Reasoning" were not like those of the INVALSI examinations. Moreover, we argued that the three dimensions used in the TIMSS examinations and the INVALSI assessments are not the same, but for the sake of comparison we grouped some categories. Since these three dimensions of classifying knowledge are not the same as the Gagne's one, we wanted to investigate which Gagne's taxonomical level is the most common in the INVALSI examinations of mathematics for grade 10 students from the school year 2010-11 to the school year 2017-18.

Our research questions were hence the following:
(1) Which Gagne's taxonomic level is the most common in the INVALSI assessments of knowledge of mathematics for grade 10 students?
(2) Which is the distribution of questions type among the Gagne's levels? Are open- and closed-type questions equally distributed among all four Gagne's taxonomic levels? Is there any taxonomic level which has a preponderant number of open or closed questions?

(3) Among the four mathematical topics, which is the preponderant Gange's taxonomic level?

Research method
In the research we used the descriptive statistical method and the non-experimental method for causal analysis.

Statistical sample
In order to have the biggest statistical sample, we decided to consider seven INVALSI assessments of mathematics for the second year of high schools with Slovene as language of instruction. The sample was composed by all the assessments from the school year 2010-11 to the year 2016-17. We decided to omit the INVALSI examination from the school year 2017-18, since it was almost identical as the assessment from the school year 2013-14.

Analysis of the Data
We investigated the Gagne's taxonomic level of each question that appears in the various INVALSI assessments of mathematics. Then we analyzed the frequency of each taxonomic level in order to understand which taxonomic level was the most popular in the INVALSI examinations. Through crosstabs we were able to compare the distribution of the four Gagne's levels among the question typologies and mathematical topics.

The collected data has been analyzed using the descriptive statistical methods, expressing the frequency of the appearance. All data was analyzed with the help of the statistical software Jamovi.

Results and Discussion
The frequencies of the topic domains
As stated in the Quadro di Riferimento (2017, 2018), the INVALSI assessments of mathematics for grade 10 students are composed by questions from the following four topic domains: "Geometry", "Numbers and quantities", "Relations and functions", and "Data and prevision". In order to understand the frequencies of each mathematical topic, we analyzed all the INVALSI assessments in the statistical sample and we labeled each question (and question item) with the main category the questions referred to. Many questions could fit in more than one category (Quadro di Riferimento, 2017, 2018), for example an exercise of geometry could require also a deep knowledge of percentages ("by how much would the area of a square increase, if we consider a 20% longer side?"), or vice versa. In this case, we considered the main topic domain the question was about. Hence, there could be some differences

between our analysis and the official data published by the INVALSI institute.

The questions in the INVALSI examinations are not equally distributed among the main four categories, as it is shown in Table 6. In Table 6 it is also clear that the most frequent category of questions is “Numbers and quantities” (36,9%), the less frequent topic of questions is “Relations and functions” (17,2%).

Table 6: The frequency of the topic domains (G = Geometry, N = Numbers and quantities, R = Relations and functions, D = Data and previsions).

Topic	Frequency (% frequency)							
	2010-11	2011-12	2012-13	2013-14	2014-15	2015-16	2016-17	Total
G	12 (29,3 %)	7 (15,6 %)	12 (27,3 %)	9 (23,7 %)	8 (19,0 %)	7 (17,5 %)	8 (20,0 %)	63 (21,7 %)
N	21 (51,2 %)	25 (55,6 %)	11 (25,0 %)	15 (39,5 %)	12 (28,6 %)	15 (37,5 %)	8 (20,0 %)	107 (36,9 %)
R	5 (12,2 %)	5 (11,1 %)	11 (25,5 %)	5 (13,2 %)	7 (16,7 %)	6 (15,0 %)	11 (27,5 %)	50 (17,2 %)
D	3 (7,3%)	8 (17,8 %)	10 (22,7 %)	9 (23,7 %)	15 (35,7 %)	12 (30,0 %)	13 (32,5 %)	70 (24,1 %)
TOT	41 (100, 0%)	45 (100, 0%)	44 (100, 0%)	38 (100, 0%)	42 (100, 0%)	40 (100, 0%)	40 (100, 0%)	290 (100, 0%)

The frequencies of the question typologies

In order to understand the frequency of the question typologies (open and closed questions), we took the data present in the work of Doz (2019). These data are replicated in Table 7. From this table we can notice that there are more closed-type questions (55,5%) than open-type (44,5%).

Table 7: The frequencies of the question typologies (C = Closed, O = Open).

Type	Frequency (% frequency)							
	2010-11	2011-12	2012-13	2013-14	2014-15	2015-16	2016-17	Total
C	26 (63,4 %)	25 (55,6 %)	30 (68,2 %)	19 (50,0 %)	20 (47,6 %)	20 (50,0 %)	21 (52,5 %)	161 (55,5 %)
O	15 (36,6 %)	20 (44,4 %)	14 (31,8 %)	19 (50,0 %)	22 (52,4 %)	20 (50,0 %)	19 (47,5 %)	129 (44,5 %)
TOT	41 (100,0 %)	45 (100,0 %)	44 (100,0 %)	38 (100,0 %)	42 (100,0 %)	40 (100,0 %)	40 (100,0 %)	290 (100,0 %)

Gagne's taxonomic levels in the INVALSI assessment

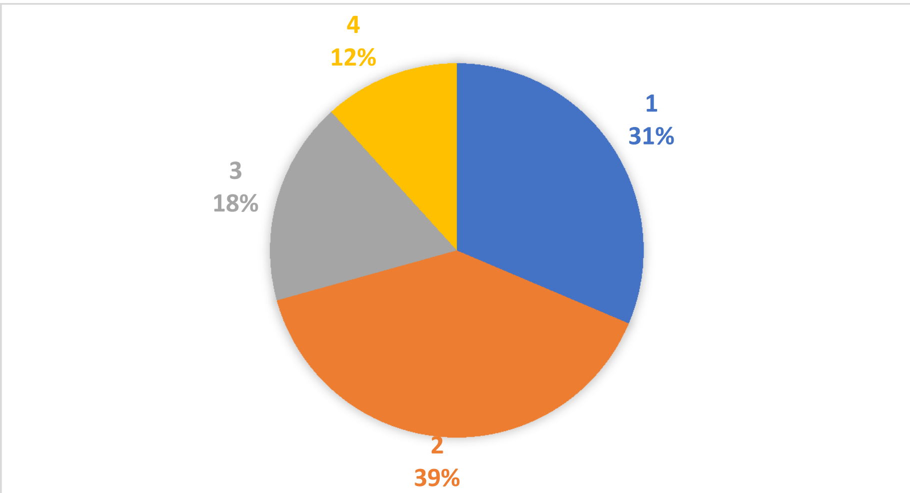
To answer our first research question, we analyzed the Gagne's taxonomic level of each question present in the INVALSI examinations in the statistical sample. We present the obtained results in Table 8 and a graphical representation is present in Figure 1. From the analysis of the single frequencies, we can notice that the most frequent Gagne's level in the INVALSI examinations, is the second (i.e. the level of routine procedural knowledge; 39,3%), followed by the first (i.e. the level of the basic and conceptual knowledge; 31,4%), the third (i.e. the level of the complex procedural knowledge; 17,6%) and, lastly, the fourth (i.e. the level of the problem-solving knowledge; 11,7%). From the analyzed data we can notice that only an 11,7% of questions are problem-solving oriented and ask the students to model a concrete, real-world situation, or to argument why a certain proof, formula, affirmation or procedure is correct.

From Table 8 we might notice that in some INVALSI assessments, the fourth Gagne's taxonomic level represented less or equal than 10,0% of all the problems in the examination (i.e. in the school years 2014-15, 2015-16, 2016-17). In the school year 2016-17, "basic and conceptual knowledge" problems represented almost half of all the exercises. "Routine procedural knowledge" represents almost half of the proposed problems in some examinations (i.e. in the 2014-15 and 2015-16 assessments).

Table 8: The frequencies of the Gagne's taxonomic levels.

Gagne's Level	Frequency (% frequency)							
	2010-11	2011-12	2012-13	2013-14	2014-15	2015-16	2016-17	Total
1	11 (26,8 %)	14 (31,1 %)	12 (27,3 %)	9 (23,7 %)	13 (31,0 %)	13 (32,5 %)	19 (47,5 %)	91 (31,4 %)
2	15 (36,6 %)	18 (40,0 %)	15 (34,1 %)	13 (34,2 %)	21 (50,0 %)	19 (47,5 %)	13 (32,5 %)	114 (39,3 %)
3	8 (19,5 %)	6 (13,3 %)	11 (25,0 %)	12 (31,6 %)	5 (11,9 %)	5 (12,5 %)	4 (10,0 %)	51 (17,6 %)
4	7 (17,1 %)	7 (15,6 %)	6 (13,6 %)	4 (10,5 %)	3 (7,1%)	3 (7,5%)	4 (10,0 %)	34 (11,7 %)
TOT	41 (100,0%)	45 (100,0%)	44 (100,0%)	38 (100,0%)	42 (100,0%)	40 (100,0%)	40 (100,0%)	290 (100,0%)

Figure 1: Gagne's Taxonomy levels in the INVALSI mathematics assessments.



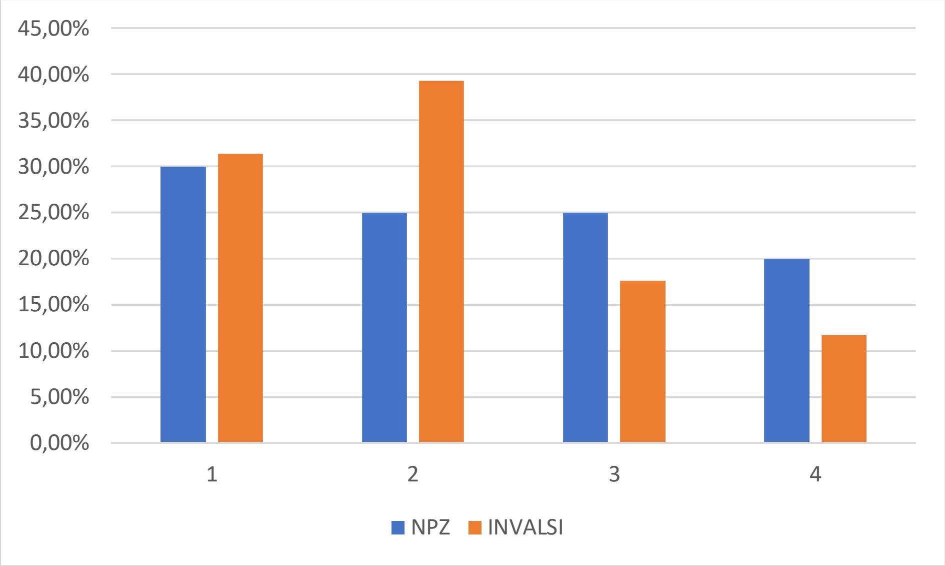
We might compare the obtained results with the results of the Slovenian national examination of knowledge (*Nacionalno preverjanje znanja, NPZ*) for grade 9 students. In the document RIC (2019) it is presented a clear classification of questions among the four Gagne's taxonomic levels. In the Slovenian NPZ of mathematics for grade 9 students, the questions are (almost) equally distributed among the four Gagne's taxonomic levels. If we join the routine procedural knowledge level and the complex procedural knowledge level, we can observe that there are 50% of problems in this macro-level. In we compare the percentages of the taxonomic levels in the NPZ with the results of the INVALSI examination of knowledge presented in Table 8, we might understand that there are some differences between the two national assessments of knowledge concerning the distribution of questions among the Gagne's levels.

We present the comparison between the NPZ and the INVALSI assessments in Table 9 and graphically in Figure 2. We can notice that there is almost no difference (1,4%) between the two assessments for what it concerns the "Basic and conceptual knowledge" level (i.e. the first level of the Gagne's taxonomy), while bigger differences are present for the second, third and fourth taxonomic levels. If we group the second and third level, we can notice that there is a 6,9% difference between the two national assessments. A greater difference is present for the "Problem-solving knowledge": in the NPZ the fourth Gagne's level represent a 20,0% of the total questions, while there are only 11,7% of problems regarding the fourth level in the INVALSI examinations.

Table 9: The Gagne's taxonomic levels in the Slovenian NPZ and Italian INVALSI.

Gagne's Level	NPZ	NPZ	INVALSI	INVALSI
1	30,0%	30,0%	31,4%	31,4%
2	25,0%	50,0%	39,3%	56,9%
3	25,0%		17,6%	
4	20,0%	20,0%	11,7%	11,7%
Total	100,0%	100,0%	100,0%	100,0%

Figure 2: Difference in taxonomic levels between the INVALSI and NPZ assessment.



While classifying the problems in the various categories of Gagne's classification of knowledge, we encountered various difficulties, since the levels are strongly connected: for example, in order to understand how to compute an area of a rectangle, the students should have the necessary knowledge about areas and multiplications.

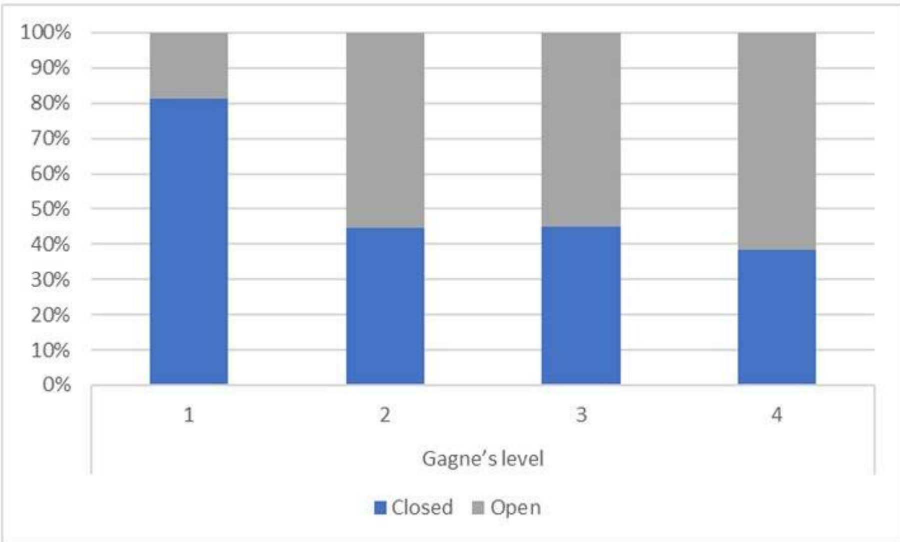
Gagne's taxonomy and question typology

To answer our second research question, regarding the distribution of the question typologies among the different Gagne's taxonomic levels, we constructed a cross table present in Table 10 (see also Figure 3). From this analysis we can deduce that the problems regarding "Basic and conceptual knowledge" are more likely closed-type (81,3%), while those regarding "Problem-solving knowledge" are more open-type (61,8%). A slight preference for open-type questions is visible for both "Procedural knowledge" taxonomic level.

Table 10: The distribution of the type of questions among different Gagne's taxonomic levels.

Type	Gagne's level				Total
	1	2	3	4	
Closed	74 (81,3%)	51 (44,7%)	23 (45,1%)	13 (38,2%)	161 (55,5%)
Open	17 (18,7%)	63 (55,3%)	28 (54,9%)	21 (61,8%)	129 (44,5%)
Total	91 (100,0%)	114 (100,0%)	51 (100,0%)	34 (100,0%)	290 (100,0%)

Figure 3: The distribution of the type of questions among Gagne's taxonomic levels.



Closed type questions are easier to correct and can evaluate students' specific knowledge of a certain topic (Powell & Gillespie, 1990). The "Basic and conceptual knowledge" could benefit the most with this typology of questions, since the knowledge of a single topic can be assessed also by true-false questions or multiple-choice one. On the other hand, closed-type questions are not suitable for problem-solving exercises, since students could use more physical space to formulate their answer, argument its validity or model a certain real-world situation. Hence open-type questions (essay-type or long-answer questions) are preferred while assessing students' abilities to formulate a hypothesis and verify its validity. Both routine and complex procedural knowledge could be assessed with both open- and closed-type questions, since once the students have solved a problem with a correct algorithm, they could write their final answer, show their work or simply choose one of the proposed options. Closed-type question do not permit, however, to check whether the students' solved the problem with the correct algorithm: the procedure could be indeed correct, but there could be some arithmetic mistakes.

Gagne's taxonomy and mathematical topics

We wanted, furthermore, to investigate how were the four Gagne's taxonomic levels distributed among the four mathematical topics.

In Table 11 we presented the distribution of the various Gagne's taxonomic levels among the various mathematical topics. As we can notice from the analyzed data, the "Routine procedural knowledge" level is the most present among all four mathematical topics, while the "Problem-solving knowledge" is the less frequent between "Geometry" (1,6%) and "Data and prevision" (4,3%) topics.

In "Numbers and quantities" and "Relations and functions", the frequency of "Problem-solving knowledge" level questions is higher than "Complex procedural knowledge", which seems to be justified by the fact that it is easier to model real-world situations that are connected to the quantities and relations between quantities (e.g. modelling the change of the water level in a tank, or modelling how would an employee's wage change in years), since students do encounter known situations. It would be, indeed, more difficult to model a real-world probability problem, since grade 10 students do not have enough (actual) experience with probability theory or statistics. Similarly, students do not have much experience with how (and why) geometry could be used to model concrete situations.

Table 11: The distribution of the Gagne's taxonomic levels among the different mathematical topics.

Topic	Gagne's level				Total
	1	2	3	4	
Geometry	20 (31,8%)	21 (33,3%)	21 (33,3%)	1 (1,6%)	63 (100,0%)
Numbers and quantities	33 (30,8%)	41 (38,3%)	14 (13,1%)	19 (17,8%)	107 (100,0%)
Relations and functions	15 (30,0%)	20 (40,0%)	4 (8,0%)	11 (22,0%)	50 (100,0%)
Data and prevision	23 (32,9%)	32 (45,7%)	12 (17,1%)	3 (4,3%)	70 (100,0%)
Total	91 (31,4%)	114 (39,3%)	51 (17,6%)	34 (11,7%)	290 (100,0%)

Conclusions

In the paper we analyzed seven INVALSI assessments of mathematics for grade 10 students. We wanted to understand which Gagne's taxonomic level was the most common in the examinations. Furthermore, we wanted to investigate, how were the questions distributed among the taxonomic level, the mathematical topic and the question typology.

We found that the most frequent taxonomic level was the "Routine procedural knowledge" (39,3%), followed by "Basic and conceptual knowledge" (31,4%), "Complex procedural knowledge" (17,6%) and, finally, by "Problem-solving knowledge" (11,7%). We compared the data with the Slovenian National examination NPZ, and we found some similarities between the two national assessments for what it concerns the first taxonomic level. Both the second and third level grouped together are more frequent in the INVALSI examination rather than in the NPZ assessment. On the other hand, the Slovenian NPZ have more exercises from the fourth Gagne's taxonomic level than the Italian INVALSI examination.

From the analysis of the frequencies of the single taxonomic levels present in the INVALSI assessments we might conclude that the Italian examination is more focused on evaluating students' knowledge and computing abilities, rather than assessing students' problem-solving abilities. Modern assessments should change their focus from evaluating merely procedures to assessing students' abilities and competences of analyzing and modeling a real-world problem (see also Cotič, 2010). From this point of view, the INVALSI assessments of mathematics for grade 10 students do not fulfil these requests, since the evaluated Gagne's taxonomic levels are mostly the first three. Hence, we might conclude that the INVALSI examinations correspond to Kulm's (1994) ideas of a "traditional tests", which are focused more on evaluating students' mechanical skills and procedures.

Our work showed that the first taxonomic level is assessed mostly with closed-type questions (81,3%), while the fourth level is assessed more likely with open-type questions (61,8%). This fact confirms the idea that assessing the fourth taxonomic level requires more open-type questions, such that students have the possibility to formulate their hypothesis, show their work, model a real-world situation, draw graphs and "discover" a formula that could described the analyzed phenomenon. There is, however, a still big number (38,2%) of closed-type questions that have the role of assessing students' "Problem-solving knowledge", which seems contradictory, since the INVALSI examinations are hence not assessing students' reasoning abilities and modelling competences.

We have also shown that the questions of each taxonomic level are differently distributed among the mathematical topics that are assessed. For instance, we found that there are relatively very few problems concerning the "Problem-solving knowledge" level among "Geometry" and "Data and prevision" problems. On the other hand, the fourth Gagne's taxonomic level is more frequent in "Numbers and quantities" and "Relations and functions". Having analyzed the content of such problem-solving exercises, we concluded that one of the possible explanations for this is that students have more experience of real-world problems that are related to percentages, wages, salaries and other physics problems, rather than with the applications of geometry and probability.

Bibliography

Altay, M. K., Yalvaç, B., & Yeltekin, E. (2017). 8th Grade Student's Skill of Connecting Mathematics to Real Life. Journal of Education and Training Studies, 5(10), 158-166.

Anderson, R. S., & Puckett, J. B. (2003). Asessing Student's Problem-Solving Assignments. New Directions for Teaching and Learning, 95, 81-87.

Bansilal, S. (2017). The difficulty level of a national assessment of Grade 9 mathematics: The case of five schools. South African Journal of Childhood Education, 7 (1), 1–8.

Cankar, G. (2008). Nacionalno preverjanje znanja kot pripomoček za izboljšanje poučevanja. Sodobna pedagogika/Posebna izdaja, 59, 130–141.

Cotič, M. (2010). Vrednotenje matematičnega znanja in objektivnost učiteljeve ocene. Pedagoška obzorja, 25(1), 39-54.

Doz, D. (2019). The National Assessment of Mathematics in High Schools with Slovene as the Teaching Language in Italy. Revija za elementarno izobraževanje, 12(2), 155-176.
Felda, D. (2018). Preverjanje matematičnega znanja. Revija za elementarno izobraževanje, 11 (2), 175–188.
Gagne, R. M. (1972). Domains of learning. Interchange, 3, 1-8.
Gagne, R. M. (1977). The conditions of learning (3rd ed.). New York: Holt, Rinehart & Winston.
Gagne, R. M. (1984). Learning outcomes and their effects: Useful categories of human performance. American Psychologist, 39(4), 377.
Gazzetta Ufficiale (2010). Retrieved from: https://www.gazzettaufficiale.it/gunewsletter/dettaglio.jsp?service=1&datagu=2010-12-14&task=dettaglio&numgu=291&redaz=010G0232&tmstp=1292405356450, 26. 5. 2019.
Guskey, T. R. (2019). Grades versus comments: Research on student feedback. Phi Delta Kappan, 101(3), 42-47.
INVALSI (2016). Retrieved from: https://www.invalsi.it/invalsi/doc_evidenza/2016/07_Rapporto_Prove_INVALSI_2016.pdf, 1. 6. 2019.
INVALSI (2017). Retrieved from: https://www.invalsi.it/invalsi/doc_eventi/2017/Rapporto_Prove_INVALSI_2017.pdf, 26. 5. 2019.
Kartianom, K., Mardapi, D. (2017). The utilization of junior high school mathematics national examination data: A conceptual error diagnosis. Research and Evaluation in Education, 3 (2), 163–173.
Kulm, G. (1994). Mathematics assessment: What works in the classroom. San Francisco, CA: Jossey Bass Inc.
Lester Jr, F. K. (2013). Thoughts about research on mathematical problem-solving instruction. The mathematics enthusiast, 10(1), 245-278.
Powell, J., Gillespie, C. (1990). Assessment: All Test Are Not Created Equally. Annual Meeting of the American Reading Forum (11th, Sarasota, FL, 12-15 December 1990), 1–13.
Quadro di Riferimento (2017). Retrieved from: http://www.invalsi.it/invalsi/doc_evidenza/2017/QdR2017_190417.pdf, 26. 5. 2019.
Quadro di Riferimento (2018). Retrieved from: https://invalsi-areaprove.cineca.it/docs/file/QdR_MATEMATICA.pdf, 30. 5. 2019.
RIC (2005). Retrieved from: https://www.ric.si/mma/izhodi%C5%A1%C4%8Da%20npz%20v%20o%C5%A1/2006070611531042/, 26. 5. 2019.

RIC (2015). Retrieved from: https://www.ric.si/mma/2015%20Struktura%20NPZ%20mat%206%20r/2014090112314222/, 26. 5. 2019.

RIC (2019). Retrieved from: https://www.ric.si/mma/2019%20Struktura%20NPZ%20mat%209%202018%2019/2018062614344950/, 26. 5. 2019.

Rittle-Johnson, B., & Schneider, M. (2015). Developing conceptual and procedural knowledge of mathematics. Oxford handbook of numerical cognition, 1118-1134.

Rosli, R., Goldsby, D., & Capraro, M. M. (2013). Assessing students' mathematical problem-solving and problem-posing skills. Asian social science, 9(16), 54.

Sulistyaningsih, E., Sugiman, S. (2016). The effect of CBT national examination policy in terms of senior high school students' cognitive readiness and anxiety facing mathematics tests in DIY Province. Jurnal Riset Pendidikan Matematika, 3 (2), 198–208.

Sumarna, N., & Herman, T. (2017, February). The increase of critical thinking skills through mathematical investigation approach. In Journal of Physics: Conference Series (Vol. 812, No. 1, p. 012067). IOP Publishing.

TIMSS (2019). Retrieved from: http://timss2019.org/wp-content/uploads/frameworks/T19-Assessment-Frameworks.pdf, 1. 6. 2019.

Thompson, T. (2008). Mathematics teachers' interpretation of higher-order thinking in Bloom's taxonomy. International electronic journal of mathematics education, 3(2), 96-109.

Vos, P. (2018). "How Real People Really Need Mathematics in the Real World"—Authenticity in Mathematics Education. Education Sciences, 8(4), 195.

Yuanita, P., Zulnaidi, H., & Zakaria, E. (2018). The effectiveness of Realistic Mathematics Education approach: The role of mathematical representation as mediator between mathematical belief and problem solving. PloS one, 13(9), e0204847.

Zhang, L., Zhang, X., Duan, Y., Fu, Z., & Wang, Y. (2010). Evaluation of Learning Performance of E-Learning in China: A Methodology Based on Change of Internal Mental Model of Learners. Turkish Online Journal of Educational Technology-TOJET, 9(1), 70-82.

Zuya, H. E. (2017). Prospective teachers' conceptual and procedural knowledge in mathematics: The case of algebra. American Journal of Educational Research, 5(3), 310-315.