Scientia in educatione, 8(2), 2017
p. 100–129
ISSN 1804-7106

Links between Success in Non-measurement and Calculation Tasks in Area and Volume Measurement and Pupils' Problems

Veronika Tůmová, Naďa Vondrová

Abstract

Measurement in geometry is one of the key areas of school mathematics, however, pupils make serious mistakes when solving problems involving measurement and hold misconceptions. This article focuses on the possible links between lower secondary pupils' ($n = 870$) success in solving non-measurement tasks and calculations tasks on area and volume and on their problems when solving measurement tasks. The study uses a mixed research design. Statistical methods are used to find correlations between the two types of tasks and a qualitative analysis is carried out to identify mistakes and misconceptions. The results show that there are indeed relatively strong links between success in non-measurement tasks and calculation tasks, and consequently, when teaching, attention must be paid to the development of both types of skills. The study identified pupils' mistakes in tasks which are within the Czech curriculum but which proved to be difficult, such as a missing link between algebraic and geometric representations, a tendency to linearize and/or to employ pseudo-analytical thinking. The study identified differences between individual classes which point to the significant role of the teacher and/or influence of the textbook used.

Key words: measurement, geometry, area, volume, structuring space, formulas.

Vztah mezi úspěšností v nenumerických a numerických úlohách na obsah a objem a žákovské problémy

Abstrakt

Míra v geometrii patří mezi klíčové oblasti školské matematiky, žáci však dělají při řešení úloh na míru závažné chyby a objevují se u nich miskoncepce. Článek se zabývá souvislostí mezi úspěšností žáků druhého stupně ($n = 870$) při řešení nenumerických a numerických úloh na obsah a objem a jejich problémy při řešení těchto úloh. Výzkum má smíšený charakter. Statistické metody jsou použity pro zjištění korelací mezi oběma typy úloh a kvalitativní analýza má za cíl identifikovat chyby a miskoncepce. Výsledky ukazují, že existuje silné spojení mezi úspěšností v nenumerických a numerických úlohách, a proto se při výuce musí dbát na rozvoj obou typů dovedností. Výzkum identifikoval žákovské chyby u úloh, které patří do českého kurikula, avšak ukázaly se jako pro žáky příliš obtížné, např. chybějící spojení mezi algebraickou a geometrickou reprezentací, tendence linearizovat a pracovat v pseudo-analytickém módu. Výzkum také identifikoval rozdíly mezi jednotlivými třídami, což ukazuje na důležitou roli učitele a použitých učebnic.

Klíčová slova: míra v geometrii, obsah, objem, strukturace prostoru, vzorce.

Introduction

Measurement in geometry, that is the notions of length, area and volume measurement and calculations, is a key area of mathematics, and not only because it has strong real-life applications. Understanding geometric measurement in school mathematics means "(1) conceptualizing that numbers can be used to quantify the amount of the attribute (e.g., length, area, volume) contained in a geometric object by determining the number of attribute-units that fit in the object and (b) being able to implement procedures for assigning measurements to objects (e.g., iterating units, using a ruler, choosing appropriate measurement units)" (Battista, 2007: p. 891).

However, pupils often do not have a good understanding of area and volume measurement and calculations (see section 1.2). In the Czech context, national and international tests point to problems pupils have with structuring space and connecting it to multiplication, and to their low ability to use other strategies than formulas, such as de-composing and re-composing (Vondrová, 2015). Rendl and Vondrová (2014) identified problems of Grade 8 pupils with the use of geometric formulas in TIMSS 2007 and attributed it to their failure to link algebraic expressions and geometric representations. The follow-up qualitative research showed, among others, the inability of average lower-secondary pupils to use a rectangular structure to find area and volume (Vondrová, 2015). Moreover, the pupils' way of tiling a shape highly correlated with their success in problems on area and volume (Tůmová & Janda, 2014) which opens a question of possible links between structuring space and solving calculation problems in measurement. Apart from single studies (Divišová, 2012; Kuřina, 2011; Vondrová, 2015), there is no systematic research aimed at measurement in the Czech context. The aim of this study is to investigate the ability of lower-secondary pupils to solve problems in geometric measurement, to bring insight into their mistakes and misconceptions and to investigate possible links between spatial skills and the ability to solve calculation tasks in measurement.

1 Theoretical framework and related literature

It is generally agreed that the concepts of length, area and volume are based on several key concepts (Curry, Mitchelmore & Outhred, 2005; Battista, 2007): conservation, unit, numerical processes and algebraic representation. When measuring area (length, volume), the assumption is that the units are conserved and can be combined in different ways (Piaget, Inhelder & Szeminska, 1960). For example, pupils come to understand that when a shape is divided into parts and these parts are re-arranged, the area remains the same. To find a number which expresses the area (length, volume), a unit is selected and iterated until the shape is fully covered. When arranging units into rows and columns, pupils come to understand that the area, for example, depends on the number of rows and the number of columns and that there is a multiplicative relationship between these numbers. The algebraic representation consists of writing a formula in numbers and later in letters.

1.1 Learning trajectory for area and volume

While most researchers concentrate on the enumeration of units as the main notion in the learning of area and volume (Outhred & Mitchelmore, 2000; Battista,

2004; Sarama & Clements, 2009), Battista in his later work proposes a learning trajectory for area and volume that contains two parallel streams: measurement and non-measurement reasoning. By the latter, he means reasoning that does not use numbers (Battista, 2007: p. 894). Figure 1 depicts the trajectory.

De-composition and re-composition of shapes is hypothesized to be crucial for understanding measurement in geometry (e.g., Rahim & Siddo, 2012), which is also supported by a historical parallel. The starting point for the development of measurement in Euclidean Geometry were comparisons between quantities (with the main strategy being the "overlapping" — *epithesis*) — i.e., non-measurement activities (Zacharos, 2006).

The measurement and non-measurement streams are supposed to work best in combination. For example, Huang and Witz's (2011) intervention study showed that the optimal curriculum for learning area was "combined". In the study, the control group did not learn anything about area, the AM group had a curriculum in which an emphasis was on the application of formulas and numerical calculations used in determining areas while the geometric operations of de-composition and re-composition were only used in the discovery of the formulas, the GM group focused on relationships between shape change and area and the geometric operations were used similarly to the AM group, and finally, the GMAM group's curriculum emphasized relationships between the two streams. The GMAM group outperformed the other groups in mathematical judgement tasks and explanation tasks, only in the calculation tasks in which a simple application of a formula was required, there was no difference between the AM and GMAM groups.

Non-measurement reasoning	**Measurement reasoning**
1. Holistic visual comparison of shapes	1. Use of numbers not connected to unit iteration
2. Visual comparison of shapes by de-composing/re-composing	2. Unit iteration and enumeration (coordinated structuring of space into arrays), includes: • Units properly located only along the edges/sides • Units properly located without overlaps or gaps • Units organized into composites (layers) – repeated addition; multiplication • Operating with other units than cubes
3. Comparison of shapes by property preserving transformations/de-compositions	
	3. Operating on numerical measurement • Structuring becomes implicit — multiplication of measures
4. Integrated measurement and non-measurement reasoning (understanding formulas for non-rectangular or composite shapes, deriving length from area, etc., relation between units)	

Connections

Figure 1: Scheme of learning trajectory for area and volume (according to Battista, 2007)

Zacharos's intervention study (2006) showed how the curriculum influences pupils' solving methods. The experimental group was learning to determine area using squares or other plane figures as units while the teaching in the control group was aimed at deriving the formula for the area of a rectangle and using this formula for area calculations. In the post test, the pupils from the control group measured the sides and used multiplication to determine the area but had problems with interpreting the results of their calculations (they lacked conceptual understanding) and made mistakes more frequently. The most common mistake was calculating the area of a figure as a multiple of two side lengths when the figure was not rectangular or regular. On the other hand, the "covering/overlapping" strategies of the experimental group led much more frequently to the correct solution and the pupils understood more often what they were measuring.

1.2 Pupils' problems in measurement

Research has pointed to problematic areas in geometric measurement. First, pupils of different ages have problems with the conservation of area (Kamii & Kysh, 2006; Tan Sisman & Aksu, 2016). Second, many problems are caused by a missing link between structuring space and numeric reasoning (Battista, 2007; Huang & Witz, 2013). Pupils do not see the underlying structure behind multiplication. Another source of problems are geometric formulas. Pupils fail when solving problems in which no numbers are given and thus, geometric formulas do not help and non-measurement reasoning must be used (e.g., Divišová, 2012; Kordaki, 2003; Kospentaris, Spyrou & Lappas, 2011; Kuřina, 2011). Pupils often resort to formulas even in such problems and their knowledge of formulas functions as an obstacle. In fact, using a formula when the problem concerns area or volume is often the first, automatic strategy pupils resort to even before trying to imagine the situation (Vondrová, 2015), which can be seen as a manifestation of pseudo-analytical thinking (Vinner, 1997). Moreover, the formula for the area of a rectangle is such a frequent and strong model of area, that it replaces the concept of area as such for some pupils and they use it for different shapes too (Herendiné-Kónya, 2015; Kamii & Kysh, 2006; Kordaki & Potari, 2002; Vondrová, 2015; Zacharos, 2006).

Pupils' misconceptions in geometric measurement include "shapes of the same perimeter have the same area" (Dembo, Levin & Siegler, 1997) and "the area (volume) grows in a liner way, the same as the side of the shape" (De Bock et al., 2007). Pupils have frequent problems with geometric units (Battista, 2007; Tan Sisman & Aksu, 2016; Vondrová, 2015). Quite a comprehensive list of errors and misconceptions of Grade 6 pupils when solving conceptually and procedurally oriented tasks is given in (Tan Sisman & Aksu, 2016). To sum up, pupils of different ages, and worldwide, seem to hold similar misconceptions and have similar problems in area and volume measurement tasks.

1.3 Spatial abilities and non-measurement reasoning

Non-measurement reasoning is clearly related to spatial skills or abilities.[1] Some authors (e.g., Tartre, 1990) see spatial skills as consisting of two main factors: spatial visualization and spatial orientation. Non-measurement reasoning is part of the former, that is, it relates to "the ability to mentally manipulate, rotate, twist or invert a pictorially presented stimulus object" (McGee, 1979), without changing or

[1] Some authors use *skills* (Tartre, Sarama, Friedman), while others use *abilities* (McGee, Pitallis).

moving the perceptual perspective of the person viewing the object (which relates to the second factor of spatial skills, i.e., spatial orientation).

Research indicates that there is a correlation between spatial skills and success in mathematics in general (e.g., see the meta-analysis by Friedman, 1995). Sarama and Clements (2009) in their meta-analysis of research claim that spatial skills directly support children's learning of specific topics, including measurement, and highlight a consistent finding that spatial skills and mathematics achievement are connected for older pupils.

Pittalis and Christou (2010) investigated a relationship between four types of reasoning in 3D geometry (representing 3D objects, spatial structuring, conceptualizing mathematical properties and measurement reasoning) and spatial abilities (spatial visualization, spatial orientation and spatial relations) of pupils in Grades 5 to 9. They showed that spatial abilities constitute a strong predictor of pupils' performance in the four types of 3D reasoning.

Hannighofer et al. (2011) focused on a relationship between Grade 3 and Grade 4 pupils' figural reasoning ability (tested by subscales for figural analogy and figural classification of the Kognitiver Fähigkeits-Test, Heller & Perleth, 2000, cited in Hannighofer et al., 2011) and measurement competencies, consisting of Instrumental knowledge and Measurement sense. The former means "having available straightforward and isolated measurement knowledge and procedures" (such as conversion of measures and their comparison) and the latter "having available knowledge about measures and units of measurement in everyday life and being able to apply all kinds of measurement knowledge in context situations" (p. 656) (such as knowledge of daily life sizes, knowing which unit of measurement belongs to an attribute, context problems about additive and multiplicative calculation with one attribute and multiple units of measurement). Figural reasoning was found to have the largest effect on the overall measurement competence and on the Instrumental knowledge and Measurement sense, in particular, as compared to the effect of grade and gender.

Apart from these studies, we did not find any research which would correlate pupils' spatial skills and geometric measurement. One of the aims of our study is to investigate this correlation.

1.4 Research questions

To sum up, authors seem to agree that for pupils' understanding of measurement in geometry and success in solving calculation problems, non-measurement reasoning and spatial skills are crucial. Our study addresses the following research questions:

RQ1: How successful will lower secondary pupils from different grades be in solving non-measurement and calculation tasks? What are their mistakes and misconceptions in tasks which proved to be the most difficult for them?

RQ2: How strong is the relation between non-measurement reasoning/spatial skills and success in solving problems on area and volume calculations?

2 Methodology

2.1 Research design

A mixed research design is used in our study to answer the research questions. Pupils were given a test with two parts (see section 2.3). The part A tasks evaluated their non-measurement reasoning (spatial skills) and the part B tasks were aimed at their ability to carry out area and volume calculations.

Two samples of pupils took part in our study. The test for Sample 2 differed from the test for Sample 1 in that the tasks in part B were assigned in random order. This was done with the aim of showing that the relative difficulty of individual tasks would remain roughly the same even though the order of the tasks changed and to check whether any correlations found would remain stable with a different sample. Part A of the test was the same for both samples.

2.2 Research sample

Sample 1 consists of 747 pupils from 40 classes (Grades 6 to 9) in 8 ordinary primary schools in Prague. The schools are representatives of a broad spectrum of socioeconomic backgrounds, they are schools for pupils from their immediate surroundings. Three of the schools had one class in each grade specialized (in mathematics or in languages). The number of pupils in the specialized classes is less than 10% of the sample.

Sample 2 consists of 123 pupils from 6 classes (Grades 7 to 9) from 2 primary schools in Prague. The sample includes 30% from classes specialized in mathematics.

The pupils from both samples were not selected in any way (entire classes participated).

2.3 Instrument

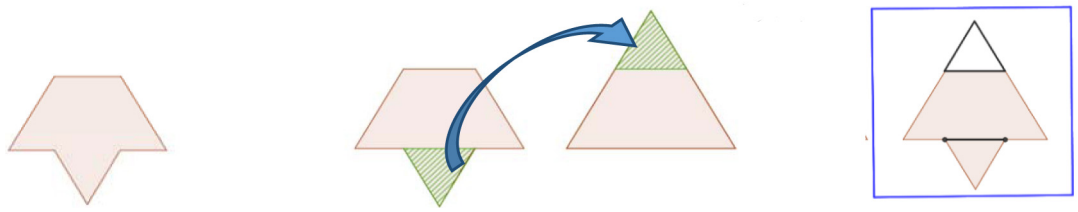
Part A of the test was to assess non-measurement reasoning related to spatial visualization skills. There are standardized (psychological) tests to assess spatial skills, however, most of them are not freely available. Therefore, we decided to modify locally developed tests already used in the Czech context (Plšková, 2010; Slezáková, 2011). Slezáková developed her test based on the "Test of Squares" which is part of Amthauer I-S-T universal intelligence tests, that are based on Rybakov figures. The sources for the test that Plšková used are not mentioned in her work, but the type of tasks corresponds to parts of the Amthauer I-S-T universal intelligence tests (2001), Differential Aptitude Test: Space Relations (Sorby, 2009) and other spatial ability tests available on-line.

Part A1 of the test finally used in our study is concerned with transformations of plane figures (de- and re-composition of shapes) and part A2 focuses on transformations in 3D space.

Part A1 consists of 20 polygons that should be modified to create an equilateral triangle (tasks U1–U20, see Appendix A). This is the assignment which included a worked-out example:

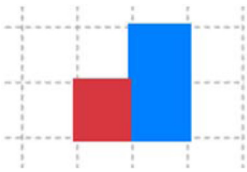
Below you will find 20 polygons, your task is to draw a single straight line where to cut the polygon so that after rearranging the parts, you get an equilateral triangle. The cut must always connect two vertices of the polygon. You also must indicate where the cut-off part should be placed as in the example below:

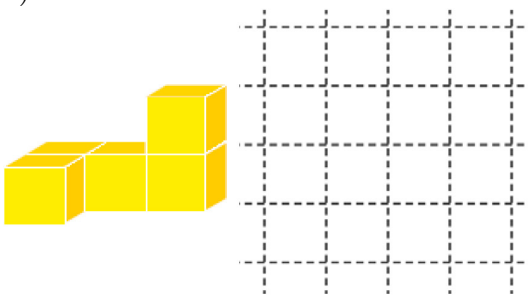
Task: You imagine: You draw:

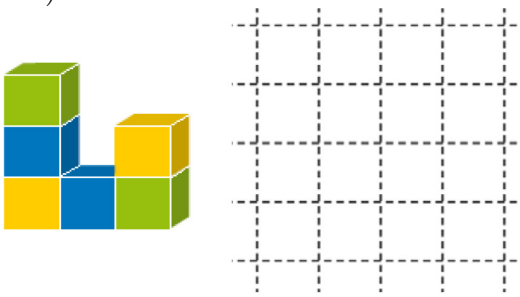


Part A2 is composed of several types of tasks that are used in testing spatial skills (Plšková, 2010; Amthauer et al., 2001, Differential Aptitude Test: Space Relations (DAT:SR) Sorby, 2009; tests available on the web[2]). There are 5 tasks in part A2 (tasks H8–H12) but some of them contain subtasks and thus this part of the test is comprised of 16 subtasks (see Appendix B). Namely, task H10 concerns mental cube rotation, tasks H9 with 4 subtasks is about mental rotation of other solids, tasks H8 and H11, each with 4 subtasks, concern the composition of a solid based on a given net, and task H12 with 3 subtasks is aimed at the calculation of the number of missing cubes in a cuboid made of small cubes. An example of a task from part A2 is H9.

Task H9: *How would the building made of cubes look from the other side (from the back)? Example:*

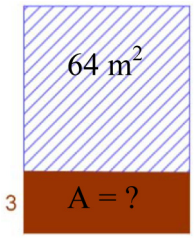
Building: You should draw: 

a) 

b) 

Part B of the test focuses on the pupils' ability to calculate areas and volumes. There are 5 tasks related to area calculation (H2–H6) and 4 tasks related to the calculation of volume (H13–H16) (see Appendix C). An attempt was made to use non-routine or novel tasks that require more advanced understanding of the underlying principles and concepts rather than a simple use of a formula. An example is H4.

Task H4: *The picture shows a plan of a garden. The hatched SQUARE represents grass and the solid rectangle is a flowerbed. Calculate the area of the flowerbed if you know that the area of grass (hatched square) is* 64 m^2 *and the shorter side of the flowerbed is* 3 m *long.*



Mathematics education researchers examined the content validity of the test and small-scale pilot testing addressed its construct validity (this is described together with a detailed a-priori analysis of all of the tasks in Tůmová, 2017b).

Note: The tasks from part A of the test will be referred to as non-measurement tasks and the tasks from part B as calculation tasks.

[2] Available from http://www.123test.com/spatial-reasoning-test/index.php; http://www.intelligencetest.com/questions/spatial.htm

2.4 Procedure

After small-scale pilot testing of the tasks and revision based on discussions with experts, the test was administered to the pupils by their mathematics teachers who had been previously instructed how to do it. Part A1 had a time constraint of 10 minutes, part A2 and B were not to exceed 45 minutes (to fit within one lesson). The test for Sample 1 was administered in September 2015, at the beginning of the school year, while the test for Sample 2 was assigned in February 2016, in the middle of the school year. Part A1 was usually done on a different day than the rest of the test, because all the parts took more than one standard lesson. For part B, the pupils were to be allowed calculators, however, some teachers decided to ban them, against our instructions. The pupils had standard equipment like pens, pencils, rulers, etc. All questions they asked during the test were recorded together with the answers from the respective teacher for us to check whether some classes did not get unintentional hints from the teacher.

The total of 747 pupils from Sample 1 completed parts A2 + B of the test, out of which 27 pupils did not complete part A1 (they were absent when it was assigned). They were excluded from the calculation of correlations between the ability to solve non-measurement tasks and calculation tasks (RQ2) but their work is used for RQ1. In Sample 2, a total of 123 pupils completed all parts of the test.

2.5 Analysis of data

The pupils' written solutions were coded by the first author and two more coders.

First, the success rates were determined. Points for non-measurement tasks were assigned in a dichotomous way: 1 for the correct answer and 0 for an incorrect one. For the calculation tasks, the scale 0, 1, 2 was used, where 2 points were assigned for the correct solution (possibly with insignificant inaccuracies or omissions), 1 point for a partially correct solution and 0 point for an incorrect solution or if the task was skipped. Aggregated results for non-measurement tasks and calculation tasks were used to assess the correlation between those two factors.

Next, several variables were defined capturing the measurement of different constructs. To measure non-measurement skills (a subset of spatial skills), the variable SPATIAL was introduced. It is calculated as a total score (expressed in a percentage of the maximum available points) in all 36 tasks and subtasks measuring spatial skills (i.e., U1–U20 and H8–H12). Based on the factor analysis of the data, the variable Ch_STRAT was defined to capture the ability to use different strategies in non-measurement tasks (this includes tasks U1, U7, U10 and U15). To measure the success of the pupils in the area and volume problems in Part B of the test, the variable CALC was defined as the total score in all nine calculation tasks (i.e., H2–H6 and H13–H16) and expressed as percentage of the maximum available points (i.e., 18). Based on the factor analysis of the results, the variable F_CONC was defined to measure success in more conceptually oriented tasks (this includes the same tasks as the variable CALC, only with tasks H2 and H3 removed). The statistical methods as depicted in section 3.2 were used to answer RQ2.

Finally, a qualitative analysis of the pupils' solutions was made to determine the pupils' mistakes and possible misconceptions (RQ1). The pupils' written solutions were coded using phenomena known from literature (e.g., the two misconceptions mentioned in section 1.2). The coding scheme was amended by new phenomena as they emerged from the written work in an open coding manner. In many cases, the

written solutions did not provide us with sufficient information to make assumptions about the possible causes of the pupils' failure and thus the numbers of mistakes of a certain type presented below must be interpreted with caution.

3 Results

3.1 Success in tasks and mistakes and misconceptions (RQ1)

3.1.1 Non-measurement tasks

Tab. 1 lists the success rates for the three most difficult tasks in each category of non-measurement tasks. For part A1 (tasks U1–U20), the order of tasks was the same for both samples, while in part A2, the order was randomized for Sample 2. The three most difficult tasks are the same in both samples.

Table 1: Success rates for the three most difficult tasks in both parts of the spatial skills test

Part A1 (2D)	Sample 1	Sample 2
U15	10,7%	22,2%
U1	15,3%	31,0%
U7	17,3%	22,2%

Part A2 (3D)	Sample 1	Sample 2
H10	28,8%	54,5%
H12.3	53,9%	71,5%
H12.1	57,0%	72,4%

The most difficult tasks in this category (for both samples) were tasks U15, U1 and U7 — see fig. 2 for the respective polygons which were to be re-arranged into an equilateral triangle. The most frequent strategy to solve tasks U1–U20 was to cut off the "smallest part that is sticking out"[3] (see fig. 3) or to place an equilateral triangle on top of the polygon and see what is missing (see fig. 4, left). However, to solve the tasks in question (U15, U1, U7), one had to apply a different strategy (see fig. 4, right) and the strategies successful in other polygons became distractors that probably prevented pupils from finding the correct solution.

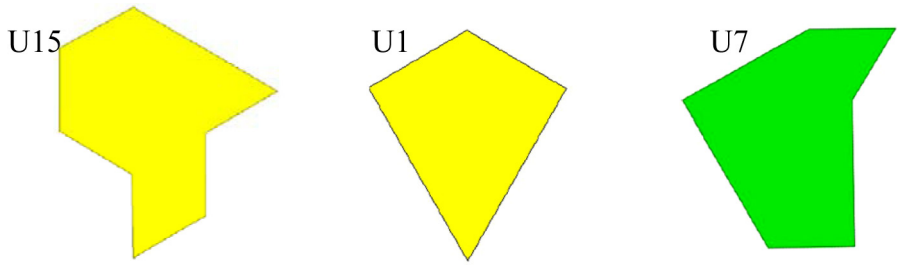


Figure 2: The most difficult tasks in the polygon de- and re-composition section (2D spatial skills test)

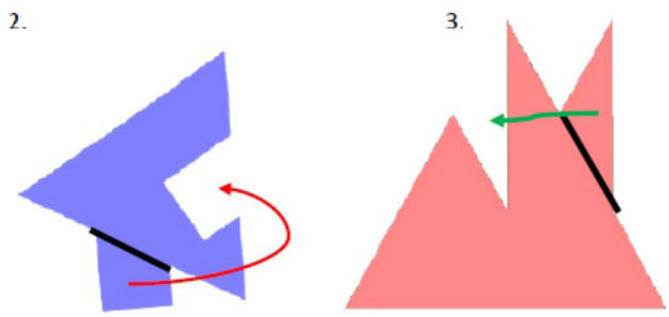


Figure 3: Strategy "cut off the smallest part that is sticking out"

[3] Which is also a strategy shown in the worked-out example.

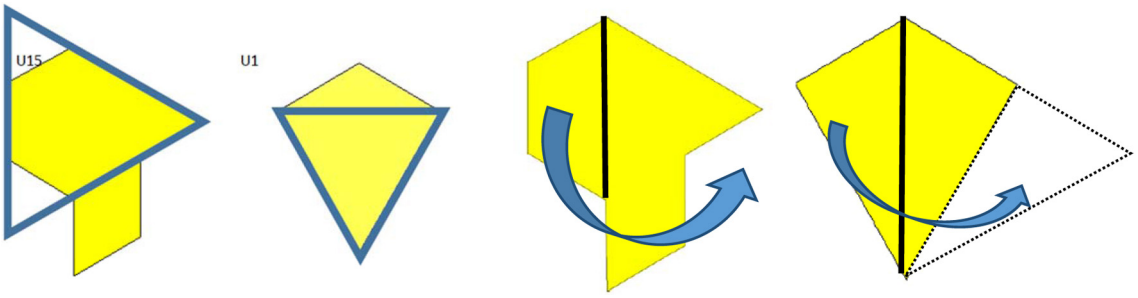


Figure 4: Incorrect (left) and correct solutions to tasks U15 and U1

In Part A2 which was dedicated to spatial skills in 3D, the most difficult task was H10 (cube rotation). The reason probably was that there is no reference object that the pupils could use for comparison. This task was followed by the subtasks of H12 (the enumeration of missing cubes in a cuboid). The most difficult of the H12 subtasks was H12.3 — a blue cube building (see fig. 5). This building has the biggest base of the three buildings (5×4) and the most frequent incorrect answer was 10 cubes (165 pupils). This is probably because these pupils calculated those cubes only that had at least one side next to the existing building and forgot to include the 2 cubes in the corners (since their sides touch the missing cubes only).

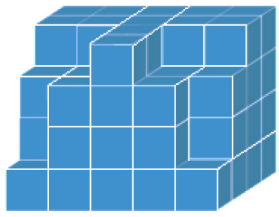


Figure 5: Blue cube building task H12.3

3.1.2 Calculation tasks

Next, we will consider Part B of the test which measured the pupils' ability to calculate areas and volumes. Before we look at the most difficult tasks and problems the pupils had with them, we will consider the grade issue — we present this analysis only for the variable CALC since the results for the variable F_CONC are similar and do not bring any new information. As could be expected, we found increasing success rates in higher grades (fig. 6). This must be interpreted with caution, however, as our study is a cross-sectional one, not a longitudinal one.

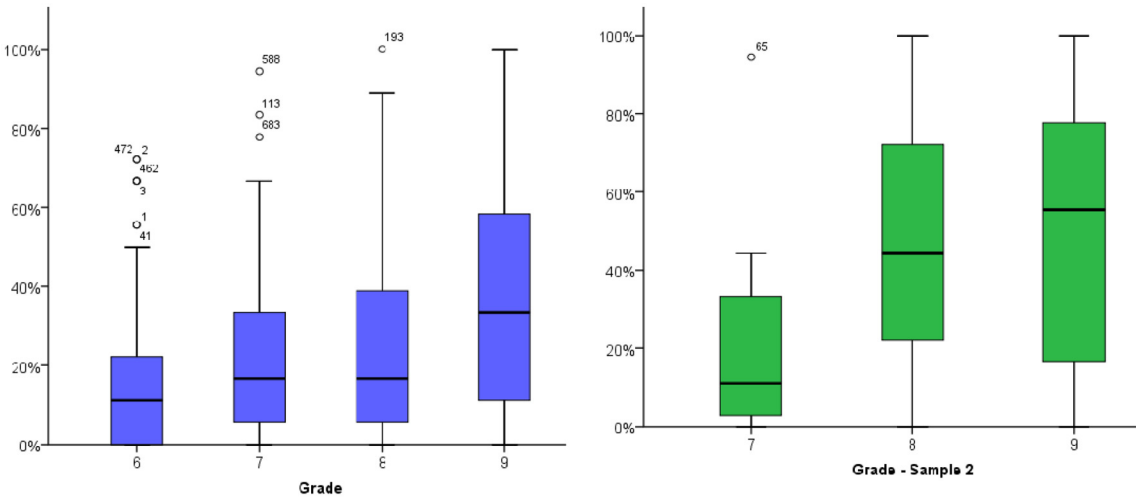


Figure 6: Box plot of CALC results per grade for Sample 1 (left) and Sample 2 (right)

Moreover, there are quite significant differences between individual classes. For Sample 1, we only show results aggregated by school and grade since showing all 40 classes would make the chart hard to read (fig. 7). For Sample 2, the two mathematics-specialized classes are in Grades 8 and 9 and are marked by hashed boxes (fig. 8).

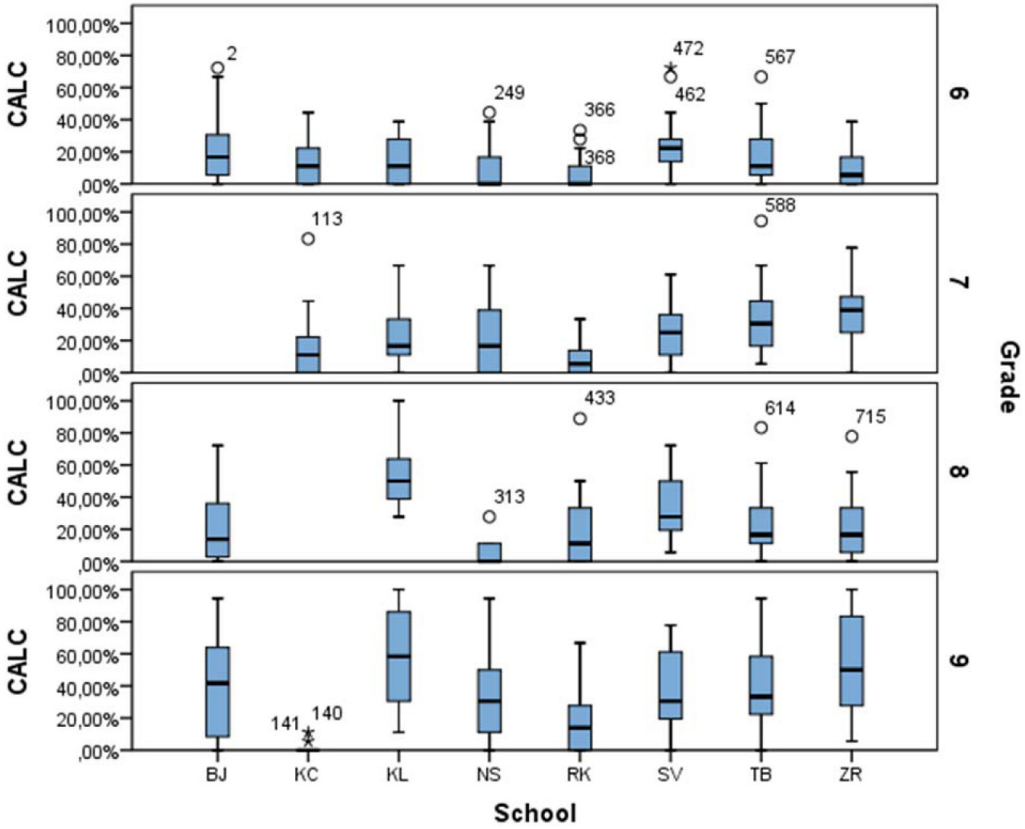


Figure 7: Sample 1 CALC results per school and grade

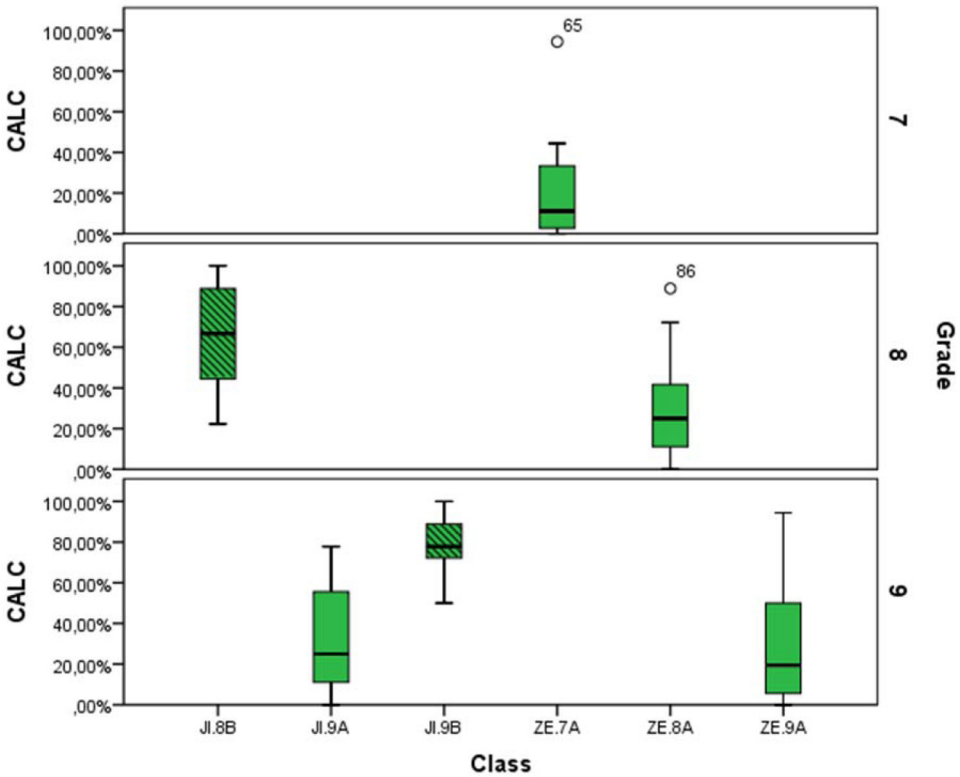


Figure 8: Sample 2 results per class

Tab. 2 shows the average success rates in the five tasks that proved to be the most difficult calculation tasks for Sample 1. The difficulty of tasks H15 and H16 is not caused only by the fact that they were the last ones in the test. These tasks were also the most difficult ones for Sample 2 where the order of tasks was random.

Table 2: Success rates for the most difficult calculation tasks for area and volume

	Sample 1 $N = 747$ (Grades 6 to 9)		Sample 2 $N = 123$ (Grades 7 to 9)	
	Success rate	No attempt	Success rate	No attempt
H15	5%	54%	22%	15%
H16	12%	74%	38%	33%
H5	15%	36%	39%	18%
H14	19%	65%	44%	35%
H6	23%	38%	38%	22%

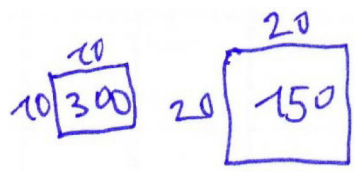


Figure 9: Solution to task H5 by a 6th grader

The most frequent misconception in tasks H5 and H15 was the assumption that if a side, or an edge, of the square, or the cubic unit, becomes 2 times bigger, so does the area, or the volume of the unit (the illusion of linearity, De Bock et al., 2007). For example, figure 9 shows a solution of a pupil from Grade 6. He assumes that if 300 of 10×10 tiles are needed, then 150 of 20×20 tiles must be used to cover the same area.

The pupils' tendency to solve the tasks based on cue words (to work in a pseudo-analytical mode, Vinner, 1997) rather than to make an image of the situation probably contributed to the small success rate. Only 10% of the pupils who attempted H5 and 5% of those attempting H15 drew a diagram to gain insight into the situation.

Task H16 required the pupils to calculate the height based on the volume and the base. Most of the pupils did not know how to solve it. 85% of Sample 1 pupils either did not attempt the task or just made a drawing or calculated the area of the base even though most of them had already learned the formula for calculating the volume of a cuboid ($l \times w \times h$). Some pupils wrote in their solution that they "lack the length of the height of the vase" which shows that they did not grasp the assignment and that when they saw that volume was involved, they wanted to use the formula to calculate it. Some solved this obstacle by making up the height which they then used in their calculations.

Task H14 (calculating the number of cuboid shaped parcels that would fit into a cubic box) motivated the pupils to use graphical solving strategies more often. The graphical or manipulative strategy usually consisted in picturing the first layer of the parcels in the box, calculating their number and multiplying it with the number of layers. The calculation strategy included determining the volume of the box and dividing it by the volume of the parcel. In both samples, almost exactly half of the pupils who tried to solve the task used a graphical strategy. Both types of strategies were equally successful — both led to the correct solution in approximately 35% of

cases (in Sample 1). The main cause for failure was either using a wrong calculation or an incorrect formula for volume (in the calculation strategy) or arranging parcels incorrectly into the first layer (graphical strategy). This points to a possible disconnection between a geometric situation and an algebraic description of the relationship (i.e., formula or operation) and to problems with structuration of space.

Task H6 is similar to a problem in TIMSS 2007 which proved to be very difficult for Czech Grade 8 pupils (their success rate was only 23%, Rendl & Vondrová, 2014). The most frequent mistake of H6 was that pupils compared perimeters rather than areas. Confusing area and perimeter is a mistake often seen in literature (e.g., Tan Sisman & Aksu, 2016). In the Czech language, it is supported by the fact that words "area" and "perimeter" start with the same letters ("obsah" vs. "obvod"). However, in H6, the word "area" does not appear so the problem must lie in conceptual understanding rather than in a simple confusion of mathematical terms. The second most frequent mistake was a wrong formula for the area of a triangle (such as $a \cdot b \cdot c$ or $a \cdot b$). Quite a number of pupils only used a visual impression of the parts of the rectangle to argue for the congruence (or not) of their areas.

3.2 Relation between non-measurement reasoning and success in calculation tasks (RQ2)

3.2.1 Variables for measurement of constructs and their characteristics

As stated above, the variable SPATIAL was introduced to measure non-measurement skills consisting of the results of 36 tasks and subtasks measuring spatial skills (i.e., U1–U20 and H8–H12). Tab. 3 shows descriptives for the variable SPATIAL for both samples. The reliability in both tests (calculated as Cronbach's alpha) is satisfactory.

Table 3: Descriptives for the variable SPATIAL for both samples

Variable SPATIAL	Mean	S.E.	Median	St. dev.	Reliability (Cronbach's α)	N
Sample 1	0.558	0.007	0.569	0.189	0.886	720
Sample 2	0.690	0.017	0.722	0.186	0.884	123

Moreover, we performed a factor analysis of the results for Sample 1 in the tasks from parts A1 and A2 to determine the main factors that constitute spatial skills as measured in our test (i.e., the variable SPATIAL). The best fitting model — we used the principal components method for factor extraction with a varimax rotation — consists of seven factors which explain more than 50% of the variability of the variable SPATIAL. The two most influential factors are correlated with tasks from part A1 — the first factor with most of the tasks U2 to U9 and the second factor with tasks U11 to U20. These factors probably describe the skills in 2D shapes decomposition and transformations and differ only in the speed with which the pupils could apply their solving strategy to different cases (tasks U11 to U20 were located on the second page of the sheet and there was a higher percentage of pupils who did not have enough time to solve them). There are four tasks in part A1 (U1, U7, U10 and U15 — the most difficult tasks in this part) that do not contribute to any of these first two factors but create a separate factor. We decided to call it "Change

of Strategy" since solving these tasks requires the ability to try various strategies in the transformation tasks (see section 3.1.1). We decided to create a separate variable out of these four tasks and call it Ch_STRAT.

The remaining four factors each correlate with subtasks of one of the tasks H8, H9, H11 and H12. It is not surprising given the fact that each of these tasks was selected from different parts of various space skills tests. Tasks H8 and H11 both work with nets of solids but their results are not very well correlated — this is probably because H8 deals with the ability to visualize the shape of a 3D solid that can be made of a certain net, while task H11 is aimed at the mental rotation of the composed solid (and the position and orientation of its sides — not only how the shape will look). Task H10 is best correlated with the factor "Change of Strategy" but the correlation is not that strong (only 0.44). The task would reduce the reliability of the composed variable, thus, it is not included in the variable Ch_STRAT. Tab. 4 shows descriptives for this variable for both samples. The reliability in both tests (calculated as Cronbach's alpha) is satisfactory (given the fact that it consists of 4 tasks only). We will use this variable in our correlation analysis as well and investigate its relationship with success in calculation tasks.

Table 4: Descriptives for the variable Ch_STRAT for both samples

Variable Ch_STRAT	Mean	S.E.	Median	St. dev.	Reliability (Cronbach's α)	N
Sample 1	0.154 9	0.007 03	0.00	0.244 5	0.607	720
Sample 2	0.272 4	0.028 60	0.25	0.316 9	0.678	123

To describe success in calculation tasks, the variable CALC was introduced as the total score in all nine calculation tasks (i.e., H2–H6 and H13–H16). Tab. 5 shows descriptives for the variable CALC in the test for both samples. The reliability (calculated as Cronbach's alpha) is satisfactory for both samples.

Table 5: Descriptives for the variable CALC

Variable CALC	Mean	S.E.	Median	St. dev.	Reliability (Cronbach's α)	N
Sample 1	0.236 6	0.008 7	0.166 7	0.237 4	0.792	747
Sample 2	0.448 1	0.029 3	0.444 4	0.324 8	0.869	123

To check whether the selected tasks measure one construct only, we performed a factor analysis on Sample 1 results of the tasks in part A2 and B to identify components of this part of the test. Tasks from part A2 were part of this analysis (but we looked at the aggregated results for the tasks H8–H12, not at results of the individual subtasks). We wanted to show that they created a different factor (i.e., the factor of spatial skills which differs from the CALC part of the test). The best interpretable model — again, we used the principal components method for factor extraction with a varimax rotation — consists of 3 factors which explain 43% of the variability of the total result. The first factor correlates well with tasks H4 to H6 and H13 to H16 and we decided to interpret it as a factor of "conceptual understanding of area and volume". The tasks included in this factor are almost identical with the tasks we designed for the variable CALC. The only difference is the exclusion of tasks H2 and H3, which are tasks not requiring a deeper understanding of the

tested concepts. Based on this factor, we decided to create the variable F_CONC that consists of 7 tasks: H4–H6 and H13–H16 and measures success in conceptually oriented area and volume tasks. Descriptives for this variable are shown in tab. 6.

Table 6: Descriptives for the variable F_CONC

Variable F_CONC	Mean	S.E.	Median	St. dev.	Reliability (Cronbach's α)	N
Sample 1	0.187 0	0.008 7	0.142 9	0.237 0	0.768	747
Sample 2	0.415 8	0.030 4	0.357 1	0.337 4	0.845	123

Note that the more difficult area (H4–H6) and volume (H14–H16) tasks are always part of one single factor in all modifications of the factor model we tried (i.e., with different numbers of factors, different extraction methods and rotations). The results in area tasks are more strongly correlated with results in volume tasks than with the easier area tasks (i.e., H2 and H3). This might suggest that the development of concepts of area and volume is parallel and connected (see also Eames et al., 2014).

Tasks H8 to H12 were indeed found to correlate well with one of the identified factors (the "space skills" factor), only the correlation of task H8 with this factor was slightly weaker (0.37). The remaining factor consisted of tasks that did not require a deeper understanding of area and volume and we decided not to use this factor in our further analyses.

3.2.2 Calculating correlations

The distribution of results in non-measurement tasks (the variable SPATIAL) and a visual check of a corresponding Q-Q plot suggest that the distribution of the variable SPATIAL might be close to normal (see fig. 10), however, Shapiro-Wilk's tests rejected the hypothesis of normality for both samples (p-value < 0.000 in both cases). The same is also true for the variable Ch_STRAT where even the Q-Q plot suggests non-normality.

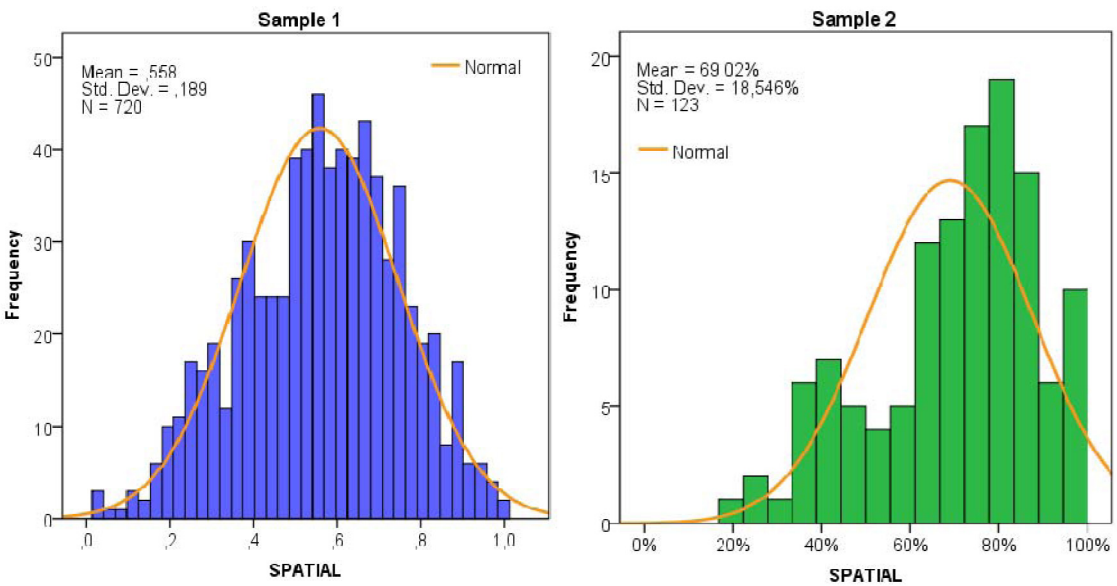


Figure 10: Distribution of the variable SPATIAL

The success rates in Part B of the test were rather low for Sample 1 (the mean for the variable CALC is less than 24%) and the number of pupils that skipped tasks in this part of the test was relatively high. The distribution of the variable CALC is, therefore, far from normal (see fig. 11, left). More than 170 Sample 1 pupils out of 747 scored 0% in these tasks. For the Sample 2 pupils, the situation was better (with a mean success rate of 44%), but the distribution of results is not normal either. This is due to the fact that two out of the six tested classes in Sample 2 are specialized in mathematics and their average scores are more than twice the results of the other regular classes which causes the "fat tail" on the right-hand side of the distribution (see fig. 11, right). The Shapiro-Wilk's tests of normality rejected the hypothesis of normality for both samples as well as for the variable F_CALC (p-value < 0.000 in all cases).

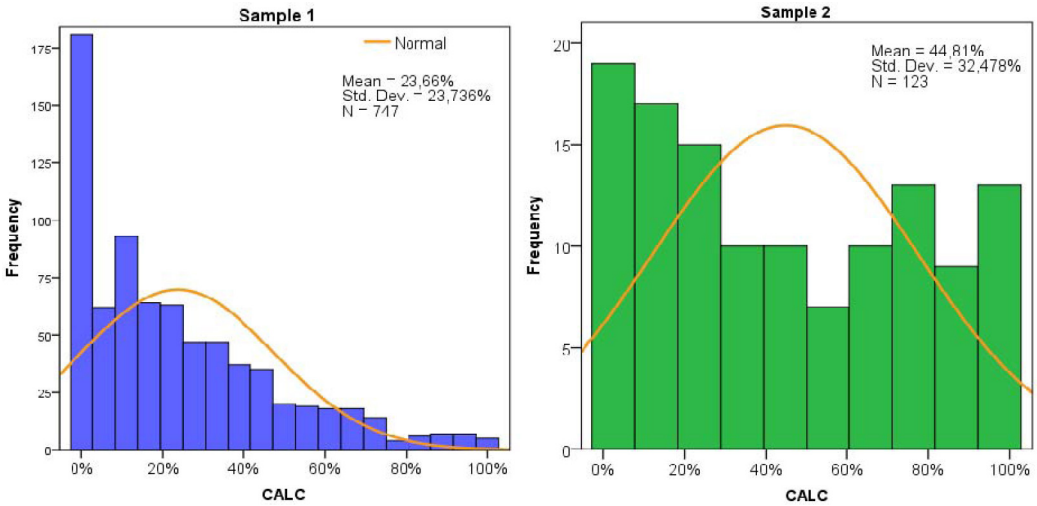


Figure 11: Distribution of the variable CALC for Sample 1 (left) and for Sample 2 (right)

Since the assumption of normality is not fulfilled for the analysed variables, we decided to use both Pearson's correlation coefficient as well as Spearman's[4] to assess the relationship between the variables measuring non-measurement reasoning (the variables SPATIAL and Ch_STRAT) and the variables measuring the success in area and volume calculations (the variable CALC for all calculation tasks and F_CONC for tasks which require conceptual understanding of area and volume). The highest correlation was found between the variables SPATIAL and CALC — for Sample 1, the Pearson's correlation is 0.591 and for Sample 2, it is 0.534. Both correlations are statistically significant (p-value < 0.000) and can be considered ranging between substantial to strong (De Vaus, 2002). The Spearman's correlations for these variables for both samples are also statistically significant (p-value < 0.000) with similar values: 0.610 for Sample 1 and 0.546 for Sample 2. Other correlations are shown in tab. 7. All the displayed correlations are statistically significant with the p-value < 0.005.

We can see that the correlations of the variable SPATIAL with success in conceptual tasks (F_CONC) are also in the range substantial to strong for both samples — with values above 0.51. The Pearson correlation of the variable Ch_STRAT (the ability to use different strategies in 2D transformation tasks) with the success in area and volume problems is less strong — only in the range moderate to substantial but still above 0.45 (De Vaus, 2002) for Sample 1. However, this correlation proved to be much weaker for Sample 2 (well below 0.3).

[4]It does not require the normality of the variables.

Table 7: Correlations between non-measurement reasoning and success in calculation tasks

Pearson correlations

	Sample 1		Sample 2	
	CALC	F_CONC	CALC	F_CONC
SPATIAL	0.591	0.531	0.534	0.513
Ch_STRAT	0.457	0.454	0.278	0.244

Spearman's rho

	Sample 1		Sample 2	
	CALC	F_CONC	CALC	F_CONC
SPATIAL	0.610	0.543	0.546	0.525
Ch_STRAT	0.361	0.353	0.284	0.251

The relatively strong relation between the variables SPATIAL and CALC is clearly apparent from the scatter plot in fig. 12. Only the data for Sample 1 are shown as the picture is similar for Sample 2. If used as a predictor, the variable SPATIAL can explain almost 35% (i.e., 0.591 squared) of the variance of the target variable CALC.

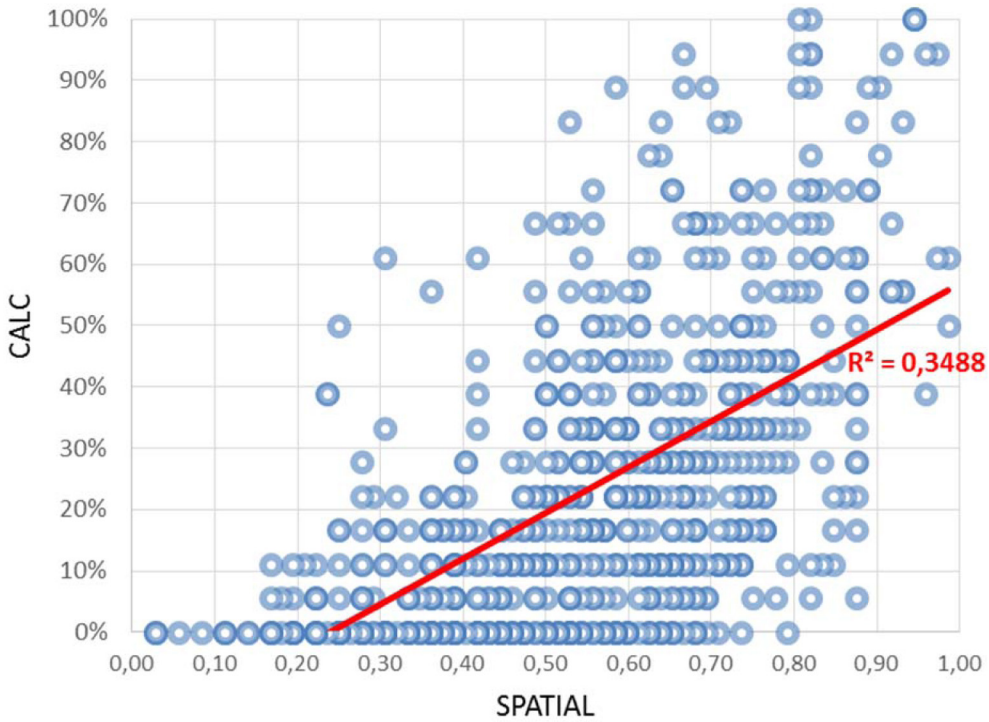


Figure 12: Scatter plot for the variables CALC and SPATIAL per pupil (Sample 1)

An even stronger relationship can be seen when we plot the average CALC result against the average SPATIAL result for each class (see fig. 13). Here, the average SPATIAL result predicts almost 73% of the average CALC result (Pearson correlation is 0.854). There are 4 classes (labelled by acronyms in fig. 13) that appear as exceptions to the trend suggested by the data. During discussion with a teacher of one of the classes (namely, ZR.9A), it came out that she pays special attention to the development of spatial skills of her pupils and that some of them were taking voluntary lessons in technical drawing. However, what we see in the graph is that this class has better CALC results than would be expected based on their average SPATIAL result. This issue might require deeper analysis.

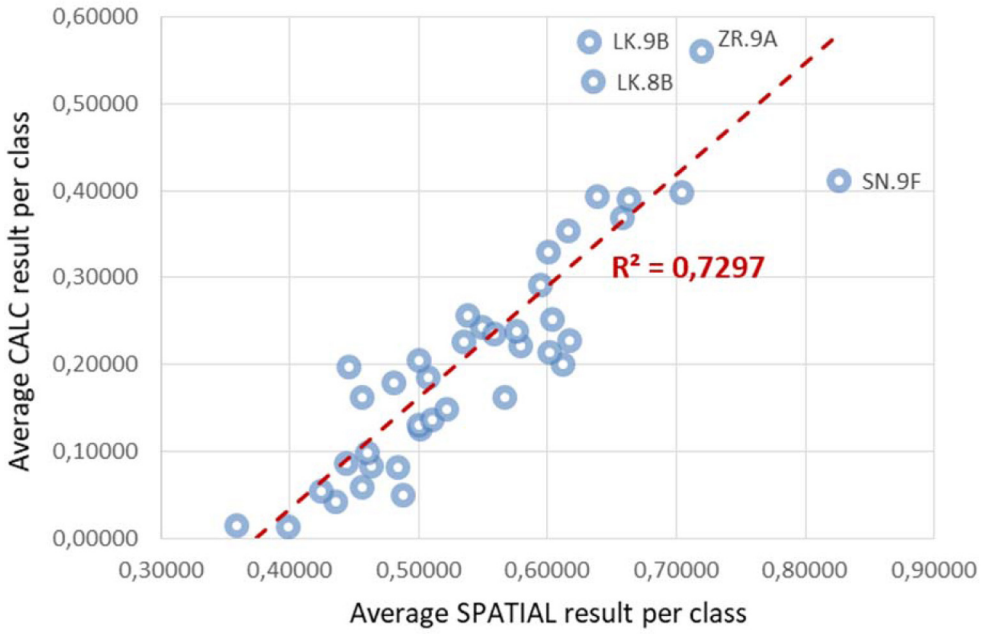


Figure 13: Scatter plot for average results for the variables CALC and SPATIAL per class (Sample 1)

4 Discussion of results

4.1 Lower-secondary pupils' success and mistakes in non-measurement and calculation tasks

Thanks to using the test in two samples, we were able to find which tasks were the most difficult ones for our pupils. The 2D spatial tasks in Part A1 seem to be more difficult than the 3D spatial tasks in Part A2. We can see two possible reasons. The first is the time restriction of Part A1, not present for Part A2. The second lies in the non-standard formulation of the tasks in Part A1 which, as our analysis shows, do not appear in the Czech mathematics textbooks used in schools.

Calculation tasks proved to be difficult for our sample even though they are tasks found within the Czech curriculum. In task H16, the pupils had to find the height based on the volume and base of the cuboid and many of them failed. Our analysis of the most used textbooks in Czech schools showed that this type of task is found in them less frequently than the task to calculate the volume based on the lengths of edges. Moreover, this problem is probably connected to difficulties Czech pupils have when working with algebraic expressions. Some of them memorize 4 different formulas for each variable instead of using only one ($V = l \times w \times h$) and deriving the appropriate variable from it. The reason might be that the formula is mostly taught earlier than rules for manipulating expressions.

Tasks H5 and H15 are among the tasks with the lowest success rate but a relatively low omission rate. The tasks probably seemed easy to the pupils, but they proved tricky. First, the pupils had to overcome a tendency to linearize which is reported to be strong in many parts of mathematics (De Bock et al., 2007). Second, the pupils might have worked in the pseudo-analytical thinking mode as described by Vinner (1997). When pupils are working in this mode, they are using superficial similarities and signals to select the fitting procedure to solve the task (without really analysing the situation — kind of finding a "quick and dirty" solution). In our case, many pupils did not make any diagram to get an insight into the task but rather used formulas and calculations based on cue words. When looking at the tasks that proved to be among the most difficult for pupils in both parts of the test,

we note that many of them are tasks for which the first-hand (or "obvious") solution does not work (tasks H5, H15, U1, U7 and U15). To solve these tasks correctly, the pupils had to "overcome" this obvious solution that did not yield the correct answer. This might be partially explained by their pseudo-analytical thinking mode. Lack of control mechanisms (e.g., checking the plausibility of the answer, analysing the situation in more depth) is typical for this type of thinking.

The above points to the fact that there might be a common underlying factor such as the culture of mathematical thinking (Hejný, 2007) that influences the pupils' success in the test. Hejný characterizes the culture of mathematical thinking as, among others, the ability to analyse the task in depth, consider various solving strategies or transfer the problem into a different modus or representation. This corresponds to our findings: in some of the most difficult non-measurement tasks (U15, U1, U7) as well as in some of the most difficult calculation tasks (H5 and H15), the pupils had to analyse the task in more depth, overcome an incorrect "obvious" solution and find a different solving strategy. It is therefore possible that pupils with a higher culture of mathematical thinking would perform better in both types of tasks (non-measurement as well as measurement) simply because they are able to take these steps. This issue calls for further investigation.

Finally, our study showed an increasing success rate between grades for calculation tasks, however, we also found differences between schools and individual classes. Leaving aside the mathematical classes, the others are not specialized and if we take into account the fact that the schools taking part in our research have similar profiles, we would not expect such a diversity. There may be many factors causing it. Among them, the influence of the teacher and the way he/she introduces geometric measurement is significant as intervention studies show (e.g., Huang, 2011; Zacharos, 2006).

4.2 Relation between non-measurement reasoning and success in area and volume calculations

When looking at the results of our study for individual pupils, their results in non-measurement tasks predict almost 35% of the variance of the results in calculation tasks and 28% of the variance for conceptual tasks. Note that the correlation remained stable for both samples which adds to the robustness of our results. If we look at average results for each class for the same two variables, the percentage of explained variance of the average CALC result (per class) grows to 73%. This, together with the fact that there are differences among classes in terms of their success in calculation tasks, points to the fact that there is a common variable related to the class (probably the teacher, his/her teaching methods, textbooks used) that strongly influences both variables. If we look at the correlations between the SPATIAL and CALC variables within individual classes, again, we see similar results as the correlation calculated for all tested pupils.

The variable Ch_STRAT that we introduced based on the factor analysis of the results and that should relate to the ability to use different strategies for solving 2D transformation tasks, was relatively strongly correlated to success in calculation tasks for Sample 1 but this was not the case for Sample 2. As stated above, the ability to use different solving strategies could indicate a higher culture of mathematical thinking (Hejný, 2007) which we expect to influence both results, calculation as well as non-measurement. Our findings are mixed since the results were not stable in both samples. One of the reasons for this might be that we selected an insuffi-

cient or incorrect indicator of the ability to change strategies (and the culture of mathematical thinking cannot be described by this one factor only). Moreover, the stronger relationship in Sample 1 might be caused by the higher number of pupils that scored 0% in Ch_STRAT (443 out of 720); 138 of them (i.e., 20% of tested pupils) also scored 0% in CALC. In Sample 2, this is not the case — only 9 pupils out of 123 scored 0% in both parts. This might indicate that the correlation in Sample 1 was over-estimated since the tasks included in both variables were simply too difficult for the tested pupils. As we concluded above, this issue would need further investigation.

As stated above, research has established relationships between spatial skills and performance in mathematics but the evidence of their link to geometric measurement is scarce. Our study results can be compared to two studies.

For 3D, Pittalis and Christou (2010) found a very strong relationship between spatial skills and measurement. They measured spatial skills including all the relevant factors (i.e., not excluding spatial orientation as in our case). To measure success in volume and area calculations, they used 4 tasks: calculation of the surface of a solid constructed by unit-sized cubes (task 1); calculation of the surface/volume of cuboids presented as open nets (tasks 2 and 3) and comparing the capacity of rectangular and cylinder reservoirs (task 4). The relationship they found might be slightly overestimated because one of the sub-tests of spatial skills (surface development) tests the ability to compose solids out of nets and this skill is directly required in tasks 2 and 3 of the calculation part. The relations we found in the present research (i.e., the relationship between spatial skills and the pupils' success in solving calculation and conceptual tasks for area and volume) are in line with these findings albeit not as strong. The reason for observing weaker correlations might be that our test measures part of the spatial skills only (and the types of tasks are not equally represented, for example, cube rotation is only used in one task) and due to using different measurement tasks: 9 measurement tasks were used in our test including area measurement (in a 2D situation and not only in a 3D one as was used by Pittalis and Christou).

Hannighofer et al. (2011) focused on pupils' measurement sense which they assessed by tasks which can be roughly compared to our calculation tasks. They reached a conclusion that from among gender, grades and figural reasoning, the latter had the strongest effect on measurement sense. They do not present what exactly they mean by figural reasoning, however, it seems to be related to spatial skills. Thus, our results for lower secondary pupils correspond to their results which relate to elementary pupils. Hanninghofer et al. suggest that an extra year of teaching measurement does not enhance pupils' success in measurement tasks as much as the quality of their figural reasoning does. Similarly, we also suggest that the improvement of pupils' non-measurement reasoning might influence their performance in calculation tasks. Another result of Hanninghofer et al.'s study opens a new way of looking at our data: in the group of pupils with a high figural reasoning ability, there was no effect of gender on the measurement achievements, while in the group with a low figural reasoning ability, the boys outperformed the girls in measurement competence.

5 Limitations, implication and future work

Our study has its limitations. The first lies in the test samples — both were convenience samples. However, the schools had similar profiles, catered for children from

their immediate surroundings, and apart from a few classes which put more stress on mathematics, there was no specialization. Moreover, the results are very similar in both test implementations (for both samples), which should contribute to their robustness. The second limitation concerns the qualitative analysis aimed at mistakes and misconceptions as it was only based on written solutions. Therefore, we have presented the data with caution, for example, not presenting the percentage of pupils committing the mistake, only its relative occurrence in the pupils' work. Even though we used a well-established distinction of non-measurement and measurement (calculation) streams in a hypothetical learning trajectory for measurement in geometry, we necessarily used several concrete tasks to describe the pupils' ability to work within the streams. The results might have been different for a different set of tasks.

Our study has some implications. The relation found between non-measurement reasoning (i.e., a subset of spatial skills) and success in solving area and volume tasks was found to be in the range substantial to strong. Based on our data, we cannot conclude about the causality of the relationship, however, similarly to others (e.g., Pittalis & Christou, 2010), it can be tentatively suggested that better spatial skills improve the pupils' skills in measurement calculations. Certainly, non-measurement reasoning proved to be tightly related to success in solving area and volume tasks and (as shown also by Huang & Witz, 2011) should play an important role in the teaching of area and volume concepts (as suggested by Battista, 2007, and others). However, it is important to emphasize that the development of non-measurement reasoning cannot replace numerical reasoning and, therefore, both of the streams should develop in parallel and be interconnected.

During the analysis of the data, a number of interesting issues came up. First, there were differences between individual classes which point to the significant role of the teacher and/or the influence of the textbook used. Next, it would be worthwhile investigating in more detail those pupils and classes that scored high in the calculation part but low in the spatial skills part (half of those with top ten differences comes from one class) and vice versa. More research is needed to explore the way measurement in geometry is taught in the class (and possibly mathematics as such) and the way pupils from the class deal with non-measurement and calculation tasks. Finally, we have presented only those mistakes and misconceptions which are related to the most difficult tasks in each part of the test but mistakes also appeared in the solutions to other tasks as well. Interviews with pupils are needed to gain deeper insight into some of them. Some interviews have already been conducted for tasks H13 and H14 with Grades 5 and 6 pupils to see how pupils of this age group structure space, what strategies they use and what the most common problems and obstacles in solving these tasks are (Tůmová, 2017a).

Acknowledgement

The paper was financially supported by the project GAČR 16-06134S Context problems as a key to the application and understanding of mathematical concepts.

Bibliography

Amthauer, R., Brocke, B., Liepmann, D. & Beauducel, A. (2001).
Intelligenz-Struktur-Test 2000 R. Göttingen: Hogrefe, 2.

Battista, M. T. (2004). Applying cognition-based assessment to elementary school students' development of understanding of area and volume measurement. *Mathematical Thinking and Learning*, 6(2), 185–204.

Battista, M. T. (2007). The development of geometric and spatial thinking. In F. K. Lester Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (843–908). Charlotte, NC: Information Age Publishing Inc.

Curry, M., Mitchelmore, M. & Outhred, L. (2006). Development of children's understanding of length, area, and volume measurement principles. In J. Novotná, H. Moraová, M. Krátká, N. Stehlíková (Eds.), *Proceedings of the 30th annual conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, 377–384). Prague: PME.

De Bock, D., Van Dooren, W., Janssens, D. & Verschaffel, L. (2007). *The illusion of linearity: From analysis to improvement.* New York, N.Y.: Springer.

De Vauss, D. (2002). *Analyzing social science data.* London: SAGE.

Dembo, Y., Levin, I. & Siegler, R. S. (1997). A comparison of the geometric reasoning of students attending Israeli ultraorthodox and mainstream schools. *Developmental Psychology*, 33, 92–103.

Divišová, B. (2012). *Geometrické úlohy řešitelné bez výpočtu* [Geometric problems solvable without calculations] [Disertační práce]. Praha: PedF UK v Praze. Školitelka N. Vondrová.

Eames, L. C., Clements, D. H. & Sarama, J. et al. (2014). Interactions among hypothetical learning trajectories for length, area, and volume measurement. In *NCTM Research Conference.* New Orleans, LA: NCTM.

Friedman, L. (1995). The space factor in mathematics: Gender differences. *Review of Educational Research*, 65, 22–50.

Hannighofer, J., Van den Heuvel-Panhuizen, M., Weirich, S. & Robitzsch, A. (2011). Revealing German primary school students' achievement in measurement. *ZDM*, 43(5), 651–665. doi: 10.1007/s11858-011-0357-y

Hejný, M. (2007). Budování matematických schémat. [Building mathematical schemes]. In A. Hošpesová, N. Stehlíková & M. Tichá (Eds.), *Cesty zdokonalování kultury vyučování matematice* (81–122). České Budějovice: Jihočeská univerzita v Českých Budějovicích.

Heller, K. A. & Perleth, Ch. (2000). *Kognitiver Fahigkeitstest für 4.–12. Klassen, Revision (KFT 4–12+R).* [Cognitive ability test for grades 4–12, revision (KFT 4–12 + R)]. Gottingen: Hogrefe.

Herendiné-Kónya, E. (2015). The level of understanding geometric measurement. In K. Krainer & N. Vondrová (Eds.), *CERME9: Proceedings of the ninth congress of the European Society for Research in Mathematics Education* (536–542). Prague: Charles University in Prague, Faculty of Education and ERME.

Huang, H.-M. E. & Witz, K. G. (2011). Developing children's conceptual understanding of area measurement: A curriculum and teaching experiment. *Learning and Instruction*, 21(1), 1–13.

Huang, H.-M. E. & Witz, K. G. (2013). Children's conceptions of area measurement and their strategies for solving area measurement problems. *Journal of Curriculum and Teaching*, 2(1), 10–26.

Kamii, C. & Kysh, J. (2006). The difficulty of "length×width": Is a square the unit of measurement? *Journal of Mathematical Behavior*, 25, 105–115.

Kordaki, M. (2003). The effect of tools of a computer microworld on students' strategies regarding the concept of conservation of area. *Educational Studies in Mathematics*, 52(2), 177–209.

Kordaki, M. & Potari, D. (2002). The effect of area measurement tools on student strategies: the role of a computer microworld. *International Journal of Computers for Mathematical Learning*, 7, 65–100.

Kospentaris, G., Spyrou, P. & Lappas, D. (2011). Exploring students' strategies in area conservation geometrical tasks. *Educational Studies in Mathematics*, 77(1), 105–127.

Kuřina, F. (2011). *Matematika a řešení úloh.* [Mathematics and problem solving]. České Budějovice: Jihočeská univerzita v ČB.

McGee, M. G. (1979). Human spatial abilities: psychometric studies and environmental, genetic, hormonal, and neurological influences. *Psychological bulletin*, 86(5), 889–918.

Outhred, L. N. & Mitchelmore, M. C. (2000). Young children's intuitive understanding of rectangular area measurement. *Journal for Research in Mathematics Education*, 31(2), 144–167.

Piaget, J., Inhelder, B. & Szeminska, A. (1960). *The child's conception of geometry.* New York: Norton.

Pittalis, M. & Christou, C. (2010). Types of reasoning in 3D geometry thinking and their relation with spatial ability. *Educational Studies in Mathematics*, 75(2), 191–212.

Plšková, Z. (2010). *Rozvoj prostorové představivosti žáků ZŠ* [The development of three-dimensional imagination of elementary school students] [Disertační práce]. Universita Palackého v Olomouci, Pedagogická fakulta. Školitelka A. Stopenová. Dostupné z http://theses.cz/id/ht91sr/

Rahim, M. H. & Siddo, R. A. (2012). High school student-teachers attempts to justify mathematical propositions utilizing spatial structuring on shape transform. *Research in Mathematical Education*, 16(2), 107–123.

Rendl, M. & Vondrová, N. (2014). Kritická místa v matematice u českých žáků na základě výsledků šetření TIMSS 2007. [Critical issues in mathematics of Czech pupils based on TIMSS 2007 results]. *Pedagogická orientace*, 24(1), 22–57.

Sarama, J. A. & Clements, D. H. (2009). *Early childhood mathematics education research: Learning trajectories for young children.* New York: Routledge.

Slezáková, J. (2011). *Geometrická představivost v rovině* [Spatial imagination in a plane] [Disertační práce]. Universita Palackého v Olomouci, Přírodovědecká fakulta. Školitel J. Molnár.
Dostupné z http://theses.cz/id/op6350/?furl=%2Fid%2Fop6350%2F;so=nx;lang=en

Sorby, S. A. (2009). Educational research in developing 3-D spatial skills for engineering students. *International Journal of Science Education*, 31(3), 459–480.

Tan Sisman, G. & Aksu, M. (2016). A Study on Sixth Grade Students' Misconceptions and Errors in Spatial Measurement: Length, Area, and Volume. *International Journal of Science and Mathematics Education*, 14(7), 1293–1319. doi: 10.1007/s10763-015-9642-5

Tartre, L. A. (1990). Spatial orientation skill and mathematical problem solving. *Journal for Research in Mathematics Education*, 216–229.

Tůmová, V. & Janda, D. (2014). Vliv používání vizualizace a matematických operací na úspěšnost žáků v úlohách týkajících se objemu a obsahu. [Influence of using visualization and mathematical operations on pupils' success in area and volume problems]. In B. Bastl, M. Lávička (Eds.), *Sborník z konference setkání učitelů matematiky všech typů a stupňů škol 2014* (225–230). Plzeň: Vydavatelský servis.

Tůmová, V. (2017a). How do pupils of the 5th and 6th grade structure space. In J. Novotná, H. Moraová (Eds.), *International Symposium Elementary Maths Teaching SEMT '17, Proceedings. Equity and diversity in elementary mathematics education* (411–421). Prague: Charles University, Faculty of Education.

Tůmová, V. (2017b). *Chápání pojmů obsah a objem u žáků základní školy* [Disertační práce]. [Conceptions of area and volume of pupils at the elementary school. Dissertation thesis.] Univerzita Karlova, Pedagogická fakulta.

Vinner, S. (1997). The pseudo-conceptual and the pseudo-analytical thought processes in mathematics learning. *Educational Studies in Mathematics*, 34(2), 97–129.

Vondrová, N. (2015). Obtíže žáků 2. stupně ve zjišťování obsahu útvarů a objemů těles. [Elementary school pupils' difficulties when finding area and volume.] In N. Vondrová, M. Rendl et al. (Eds.), *Kritická místa matematiky základní školy v řešeních žáků* (253–318). Praha: Karolinum.

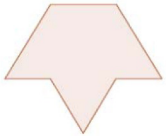
Zacharos, K. (2006). Prevailing educational practices for area measurement and students' failure in measuring areas. *The Journal of Mathematical Behaviour*, 25(3), 224–239.

Veronika Tůmová, veronika.tumova@pedf.cuni.cz
Naďa Vondrová, nada.vondrova@pedf.cuni.cz
Univerzita Karlova, Pedagogická fakulta
Magdalény Rettigové 4, 116 39 Praha 1

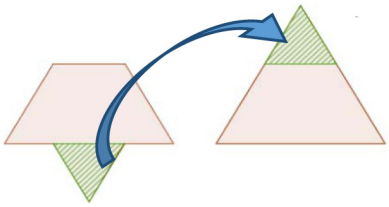
Appendix A — 2D spatial ability test

Below you will find 20 polygons, your task is to draw a single straight line where to cut the polygon so that after rearranging the parts, you get an equilateral triangle. The line for cutting must always connect two vertices of the polygon. You also must indicate where the cut-off part should be placed as in the example below:

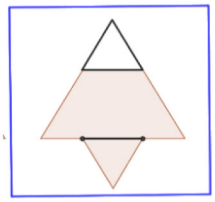
Task:



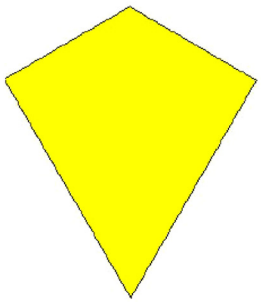
You imagine:



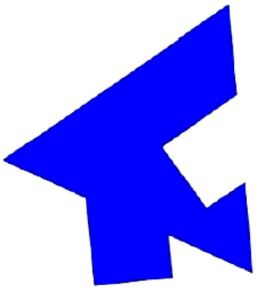
You draw:



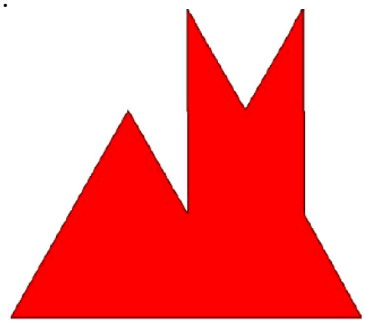
1.



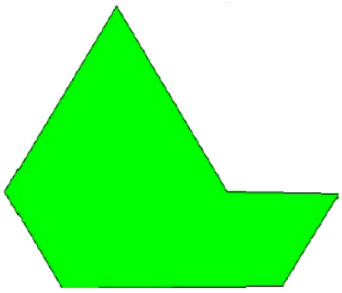
2.



3.

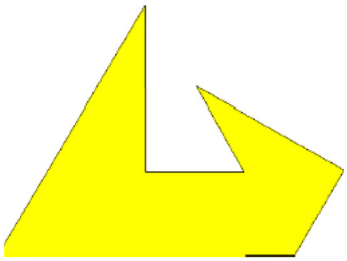


4.

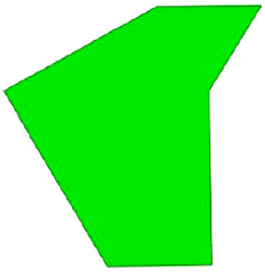


5.

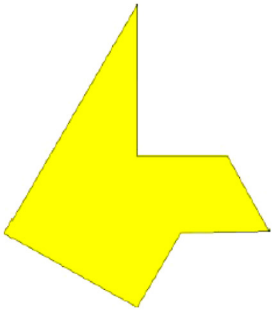
6.



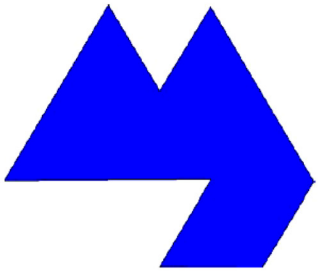
7.



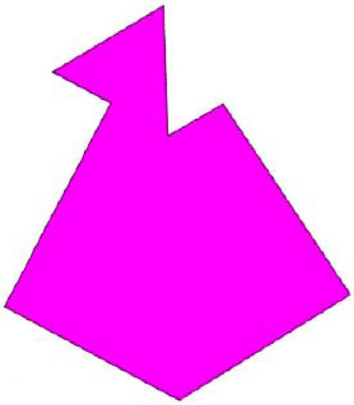
8.



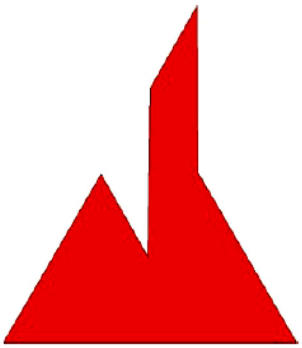
9.



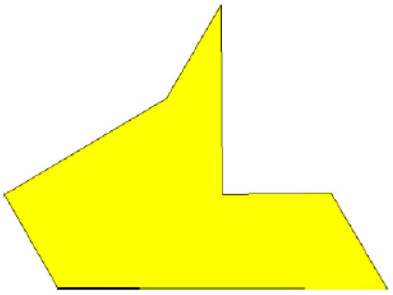
10.



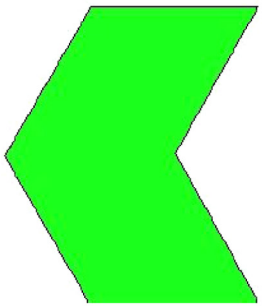
11.



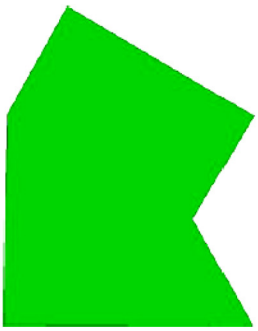
12.



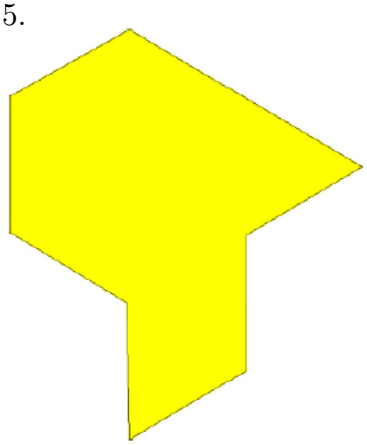
13.



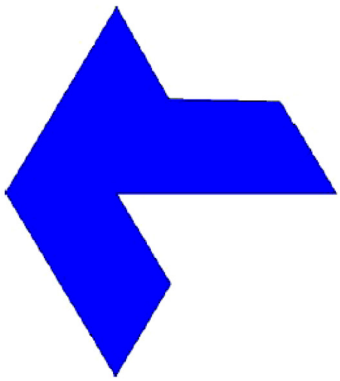
14.



15.



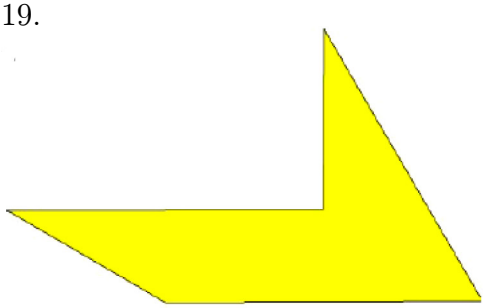
16.



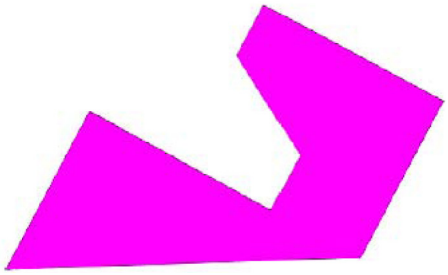
17.

18.

19.

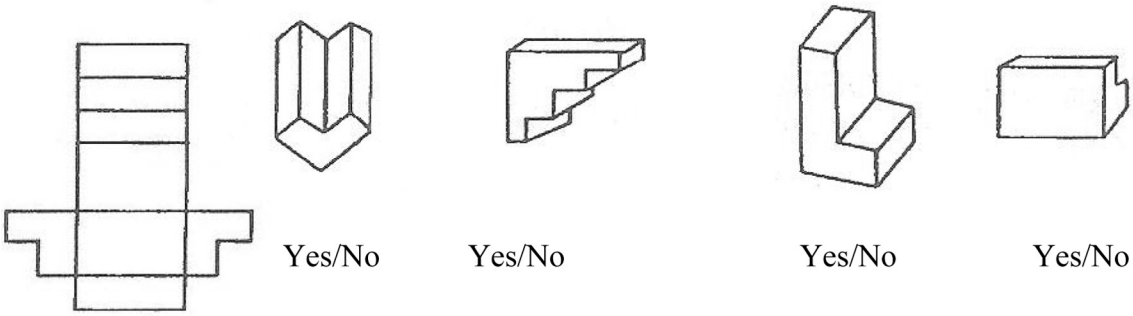


20.



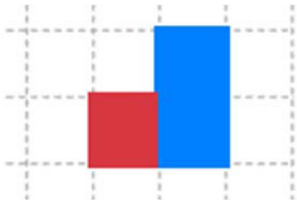
Appendix B — 3D spatial ability test

Task H8: *The picture on the left shows a template to make a box. Which of the boxes can be made using this template? Mark Yes or No for each.*

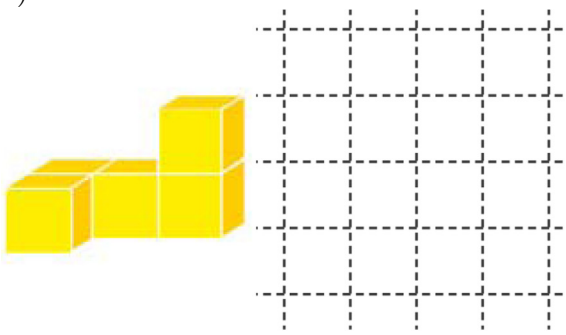


Task H9: *How would the building made of cubes look from the other side (from the back)? Example:*

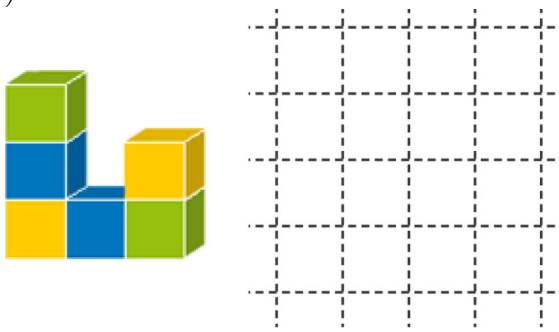
Building:

You should draw: 

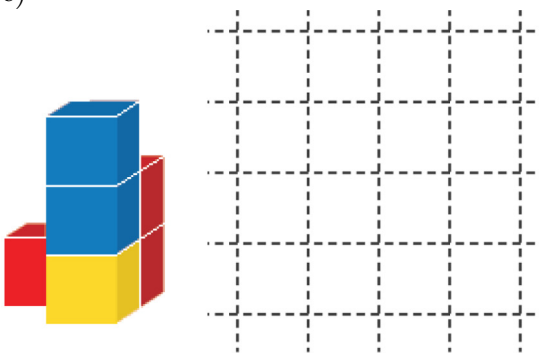
a)



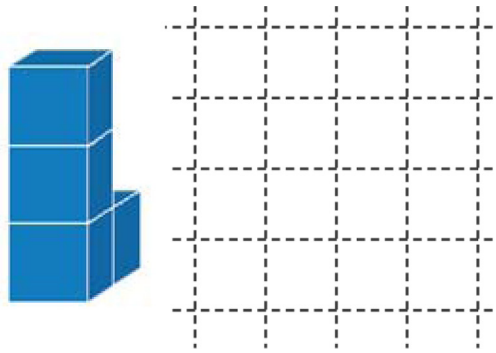
b)



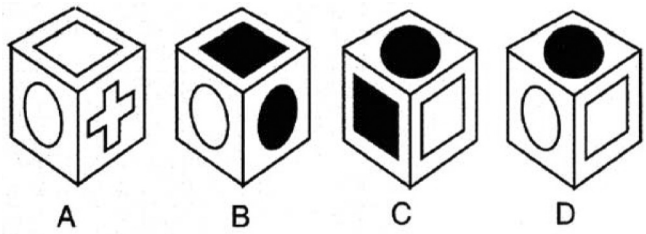
c)



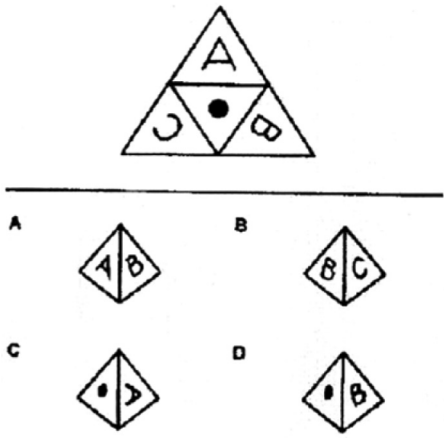
d)



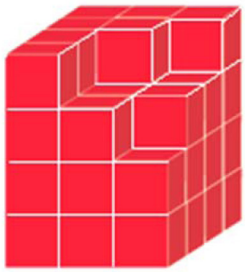
Task H10: *The pictures below show the same cube from different perspectives, however, one picture does not fit with the rest. Decide which one it is and cross it out.*

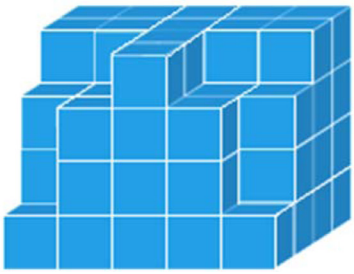


Task H11: *Which of the tetrahedra cannot be obtained from the net by folding? Cross it out.*



Task H12: *The green cube building has 9 cubes in the first layer, the red one 12 and the blue one 20. How many cubes do you have to add to get the smallest possible completely filled prism (in other words, if the building was in a tightfitting rectangular box, how many cubes would you have to add in order to fill that box completely)?*

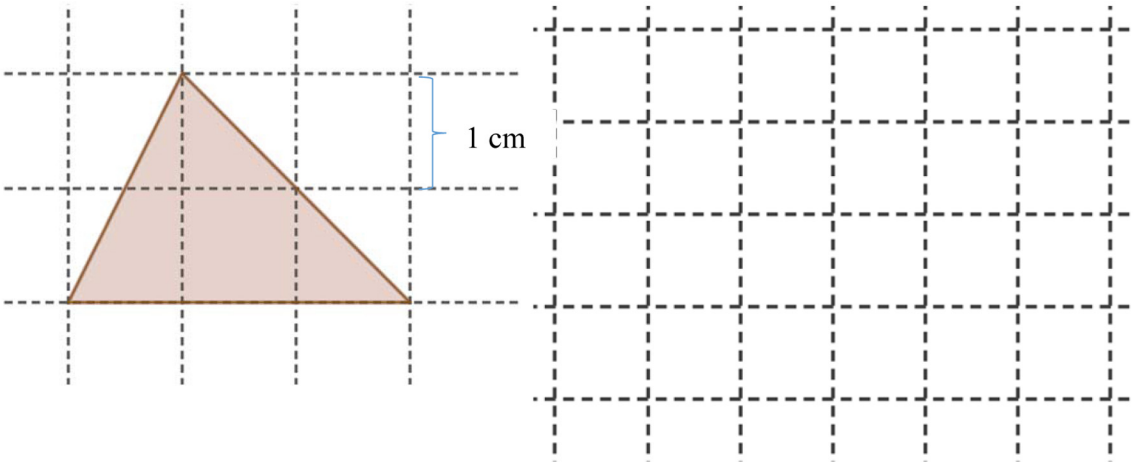




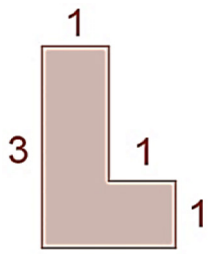
Green: ..... Red: ..... Blue: .....

Appendix C — Calculation tasks for area and volume

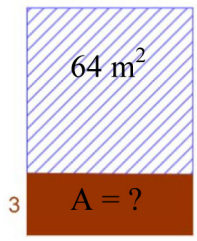
Task H2: *Determine the area of the triangle if you know that the smallest square in the grid has the side of* 1 cm. *Draw two more figures with the same area.*



Task H3: *Determine the area and perimeter of the figure in the picture.*

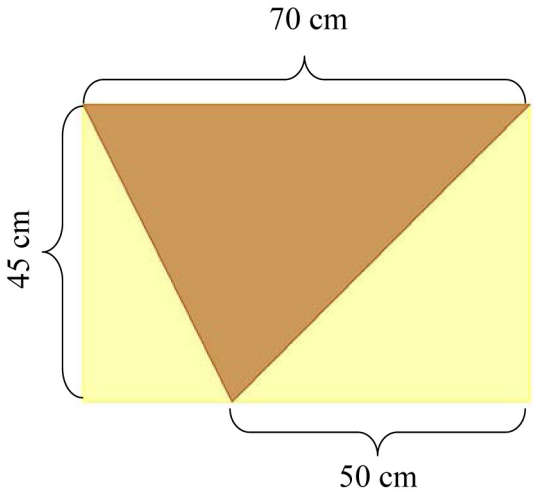


Task H4: *The picture shows a plan of a garden. The hatched SQUARE represents grass and the solid rectangle is a flowerbed. Calculate the area of the flowerbed if you know that the area of grass (hatched square) is* $64\,\mathrm{m}^2$ *and the shorter side of the flowerbed is* 3 m *long.*



Task H5: *We used 300* $10 \times 10\,\mathrm{cm}$ *tiles to pave our bathroom. Now we want to redecorate it and the new tiles are* $20 \times 20\,\mathrm{cm}$. *How many new tiles will I need? Neglect the gaps between the tiles — as if there were none.*

Task H6: *Will cut out the darker part of the rectangular board (see the picture). His friend Paul is angry with him saying: "This way, you will throw away more than half of the board." Will disagrees and claims that the rest is smaller than the part which he will use. Who is right and why?*



Task H13: *You have exactly 59 cubes (with the edge of 1 unit) to create a building on a rectangular plot of land which is 4 units long and 3 units wide. You must use all the cubes but the building has to be as LOW as possible. How many layers will there be? How many cubes will there be in the top layer?*

Task H14: *What is the maximum number of parcels measuring* $2 \times 1 \times 1\,\mathrm{dm}$ *that would fit into a cubic box with an edge of* $6\,\mathrm{dm}$*? Justify your answer.*

Task H15: *A hundred cubes (each with an edge of* $12\,\mathrm{cm}$*) fit into a box and fill it completely. How many cubes with an edge of* $4\,\mathrm{cm}$ *do you need to fill the same box completely? Justify.*

Task H16: *A cuboid-shaped vase has a base of* $9 \times 12\,\mathrm{cm}$*. If I pour one litre of water inside, how high will it reach? (Hint:* $1\,\mathrm{l} = 1\,000\,\mathrm{cm}^3$*).*