Liet. matem. rink., T. 42, spec. nr., 391–396

The validation of mathematical skills application tasks

Danutė KISELIOVA, Arkadijus KISELIOVAS, Ingrida DONIELIENĖ,
Orinta KISELIOVA (ŠU)
e-mail: kadia@su.lt

Introduction

From the point of view of classic didactics, achievements in learning mathematics is a certain system of knowledge and abilities. From the point of view of free education, achievements in learning are defined by application and critical thinking categories. From the point of view of L. Vygotsky school representatives, achievements in learning is development of higher psychic functions and for mation of scientific concepts. According to the didactics, based on J. Piaget theory, achievements in learning is the higher content that fills in the forms of thinking and gives the pupil competence in a particular field. These views in a way determine the abilities defined by Lithuanian mathematical education standards [2] though their conceptual origins are not specified.

The old point of view that the essence of mathematics is practice of calculation, is being changed by the new tendency, where mathematics is defined as school of thinking, a device of speaking and the space of solving real problems.

Construction of didactic (achievements) tests is a specific field of scientific cognition realized following strict standards [10]. In the test theory entire mathematical tools have been created meant to justify test construction procedures [2, 9, 11, 12].

Application of mathematics is urgent especially for the modern society, when new interpretations of one or another phenomenon originate, when the world becomes more complicated; the approach of the content to the real environment and to the experience clearly defines the aim of studying of some subjects.

The application of mathematic direction task performance in primary forms of Lithuanian mainstream schools traditions is not sufficiently expressed. The content of this direction tasks is determined by general education general programmes [8] and general standards [3].

The object of the investigation – mathematical skills of the fourth-form pupils.

The aim of the investigation – to select the mathematical application tasks, suitable for the fourth-formers' mathematical skills diagnostics.

The tasks of the investigation – to do psychometric mathematical applications sampling of subtest tasks, giving the examples of tasks and psychometric characteristics (coefficient of difficulty, resolution, informativeness, etc.).

The structure of test content and proportions of structural parts are established according to the standards and requirements of the general programmes of primary school for mathematics.

The calculations have been conducted using specialized statistical computer programmes packages PAULA [13] and SPSS [4].

After the initial analysis of the tasks a conclusion was drawn, that the set of mathematic application skills task test is compiled of six task groups (Table 1).

Such content of tasks absolutely satisfies the general programme requirements of primary school. The greatest part of application tasks consists of buying and choosing different purchase tasks. Buying tasks is one of the most real everyday application arithmetic operations spheres. We can see, that pupils do money calculations easier. Measure but not calculation mistakes are more frequent when doing them. Furthermore, buying is a sphere, in which the fourth-form pupils have enough experience, because they do it practically every day.

Another group of tasks which takes over 17% of all set tasks are time tasks. The results of these task solutions show the ability of pupils to find the disparity of time, to applicate the measures of time, to differentiate the concept "earlier – later". Skills in calculating time and measuring (as purchased) is the most important thing for students' everyday life.

Way tasks – are not dissociated of time. Movement as other life spheres that have been mentioned above, is common. Solution of these tasks shows pupils' ability to place objects in space and time.

30 tasks were chosen for constructing the test [7, p. 28–33], 10 – for every level; some of the tasks were made more concrete by variants (introducing additional conditions). The first 10 tasks satisfy minimal level, tasks 11–20 are for the basic and the last 10 (21–30) are for the higher level. The matrix of statistic analysis was composed of evaluations of 30 tasks (44 variants).

The sample. The investigation was conducted in March–April 2001. 361 pupil of the fourth-form did the task, 44% of them – girls, 56% of them – boys. 57% of the pupils

Table 1

Distribution of application tasks

	Components of the content	Task number	Number of tasks	Number of variants
1.	Tasks of choosing different purchases	1, 4, 11, 12, 16, 23, 24, 26, 27, 28	10	10
2.	Tasks of time calculation	5, 6, 7, 10, 14	5	5
3.	Tasks of way finding and related calculations	13, 17, 20, 21, 25	5	5
4.	Tasks of choice and action argumentation	9, 15, 19, 21, 22, 24	7	14
5.	Tasks of square measurement of geometric figures	8, 18, 29	2	2
6.	Tasks related with pupils' mode of life (school) and everyday activities	3, 30	2	2

Table 2

The results of application tasks performance

	Weak	Satisfactory	Sufficient	Good	The number of pupils
Entire sample	12.2%	14.4%	34.6%	38.8%	361
Girls	7.6%	13.2%	37.1%	42.1%	159
Boys	15.8%	15.4%	32.6%	36.1%	202

Table 3

The dependence of the application tasks performance on the textbooks

Textbook\Test	Poor	Satisfactory	Sufficient	Good	The number of pupils
„Skaičių šalis"	19.8%	12.6%	33.3%	34.3%	207
„Matematikos pasaulyje"	2.0%	16.9%	36.3%	44.8%	154

studied from the textbook „Skaičių šalis" [1], 43% – from the textbook „Matematikos pasaulyje" [6]. The results of testing are presented in Table 2.

It is not very easy to construct application tasks for the fourth formers; the difference between girls and boys is statistically significant ($x^2 = 7.94, p < 0.05$). A very great disparity we can see among the pupils who used different textbooks ($x^2 = 27.0, p < 0.0001$); only 2% of the pupils who studied from the textbook "Matematikos pasaulyje" were not able to demonstrate sufficient results in solving application tasks, and among those who used the textbook „Skaičių šalis" there were 19.8%. Such results could be determined that the content of the mathematics textbook „Skaičių šalis" is more oriented to the contention of arithmetic operation skills. The application tasks are of standard formulation asking to know the algorithms of solutions and rarely demand solutions of standard situations, there are fewer logic tasks. The relation of testing results with the textbooks used is shown in Table 3.

The statistical analysis of the tasks

The alpha factorial analysis verified that the application tasks are not homogenous; they can be dispensed into three groups. The first factor which explains the greatest amount of the dispersion of evaluations (34.4%), is related with all the tasks of minimal level and with half of basic level ones. We have to pay attention, that little factorial dimension have those tasks which were not chosen by the greatest part of pupils (num. 2. 6. 13). We shall analyse some of them.

Task 2 – *What part of a rope we'll cut if we bend it in half and that part we will bend in half again and then cut it?* – it could be, that the answer in a common fraction stopped pupils from doing this task. Even one third of all the pupils wrote the answer in how many parts they had divided the rope, they wrote the length of the designated part in centimetres and they wrote other, incorrect common fractions. The mistakes of this

tasks solution show insufficient skills of the forth-formers to link a part of a thing with a common fraction and it shows the formal presenting of this theme at school.

Task 6 – *The forth-formers have to cycle to the camp 40 kilometres. The pupils already cycle for two hours at a speed of 9 km/h. The rest part of the way the cycled in 8 km/h. How long did it take them to cycle the rest part of the way?* – the solution was more difficult, because the period in the answer is not a whole number (hours). Pupils have to know how to transform the rest of the division into minutes or to write the answer in a decimal fraction. 30.5% of pupils gave right solutions, 29% of them wrote the answer in hours and minutes, 12.2% – in a decimal fraction. It was easier to solve the task for a part of the pupils, when the teacher made changes in the task: corrected the length of the way (not 40 km, but 50 km) or the speed (not 9 km/h, but 10 km/h, or not 8 km/h, but 5 km/h). So the pupils got an integer number in the answer.

The simplicity of the problem can be the distinctive sign of that group of tasks. The next factor, which explains 22.6% of the variance of the estimate, is to be related with the rest basic level tasks and with majority of tasks of the higher level; it means that lots of these level tasks are not so indiscrete. The performance of tasks 16–27 shows only the ability to do arithmetical operations, but it shows more clearly the ability to draw conclusions in problematic situations. The majority of tasks demand to find not only the unknown quantity or to choose better solution, but to ground one's decision or draw appropriate conclusions concerning the answer. Some of the tasks (19, 21) demand to explain the solution not in words but in geometric drafts. Three last tasks are ascribable to the third group, which evaluations are determined not only by the first (simplicity) factor, which explains 7.2% of the variance of the estimate; the peculiarity of these tasks is that an extra variable has to be introduced. It is easy to do such tasks by making equations, but it has to be thoroughly considered if one solves by a traditional arithmetic way. We can say that this factor demonstrates a search of a logic scheme.

According to the difficulty of the test tasks we can group them into simple (the difficulty coefficient is 0.20 and less), of intermediate difficulty (coefficient is from 0.21 to 0.40) and difficult (coefficient 0.41 and more) ones. On the basis of this grouping we got such a distribution: 26% are simple, over 37% of intermediate difficulty and difficult. Pupils solve the problems easily, which demand only one directly applied counting operation. The problems which demand some operations are more difficult for pupils; the most difficult among the others are tasks 6, 8, 13, 20, 25, 32.

Task 13 – *The car goes 60 km/h, it uses 7l of petrol for 100 km, and when it goes 90 km/h it uses 8l for 100 km. How many litres of petrol has the car used if it covered 200 km in 60 km/h, and the other 500 km – in 90 km/h?* – a lot of pupils (31%) did not do it, because they thought that it was too difficult (5 arithmetic operations has to be done); but this difficulty was only supposed, because for the solving of this problem you do not need particular deduction, a mode of life thinking level is enough.

The most difficult was the 27th task: *It was paid as much for 2 metres of a silk band as for 3 metres of a linen band. All purchase cost 24 Lt. How much is 1 metre of a silk band more expensive than 1 metre of linen band?* 45% of pupils made mistakes in this task, 31% did not do it at all, it could be that pupils didn't understand that the price of both

bands made up only half of the price of all the purchase (usually such a task is expressed *equally*).

The minimal coefficient of informativeness is equal to 0.36% of the main evaluation scale (3.3 variant), maximum – 5.05% (24.1 variant); medium informativeness is 2.63% of scale's total mark; on the basis of this, the tasks of low, medium and high informativeness were distinguished. The relation of difficulty and informativeness when doing these tasks, shows Table 4.

The relation between difficulty and informativeness is clear – the coefficient of contingention is $C = 0.53$ ($p < 0.01$); easy variants are not of high informativeness, 34% – of high informativeness; the other variants are of medium informativeness. Some of the tasks will be discussed in detail.

Task 9 – *Continue the string of numbers 90, 180, 270, . . .; 955, 905, 855,. . .* – for the pupils of the forth-form was very easy, most of them did it correctly (though there were such pupils of strong groups who did not do the task).

Task 10. *The wheel turns around 52 times per 4 minutes. How long does it take to turn around per hour?* 30% of pupils did not do this task. The solution of the task is estimated in 14 points by a thousand points scale. The pupils of lower than medium abilities (except the weakest) did this task not worse than the strongest, the worst of all did the pupils of medium abilities. A very similar result was received when analysing the execution of task 12.

The result of the last stage is statistic characteristics of measurement tasks which were left for the next stage of analysis. Three factors of tasks are left after the choice, it means that the homogeneity of tasks system is not high, but the 42–44 tasks dependence on two factors disappeared. The main factor determines 30.1% of the variance of the estimate, it takes over the tasks of the main and higher level. The interpretation of the other two factors does not change, but the set for the first factor has become less, eliminating uninformative tasks of minimal level. The reliability of the general test scale is high enough; the dividing in two halves coefficient of correlation is 0.86. The coefficient of Gutman's reliability is 0.92.

Item characteristic functions [5] of all the tasks are near to an ideal model, it means that the pupils of low ability level do them only by chance, and those who are of high abilities – do them rather often, only some of them make mistakes.

The informative tasks, initiate problems where you have to make definite operations if you want to do them, and the logic of solution is not so difficult. We can see that

Table 4

The relation of task variants difficulty and informativeness

Difficulty\Informativeness	Low	Medium	High	Total
Easy variants of tasks	8	2	–	10
Variants of medium difficulty	5	3	6	14
Variants of difficult tasks	2	6	6	14
Total	8	31	13	38

these tasks are standard; more complicated than these are less difficult tasks, aren't so informative (but they are reasonable for the test).

The chosen tasks distribute like this: various tasks of sale and purchase – 6, time calculation – 1, way finding and trips calculation tasks – 5, activity and choice argumentation – 6, square calculation of geometric figures – 1, tasks related to the real pupils' life – 2.

Conclusions

Generalizing application tasks, we can argue, that the greatest part of pupils understand not only the tasks which they read by themselves, but they are able to solve them adjusting two or more standard operations or procedures. Doing non-standard practical tasks they orientate on the strategy of solution, they are able to argue their actions. It is very important to teach pupils to do sums purposely and expediently, and not only imitating the previous resolution of tasks, reposing upon analogy. Of course such analogies are very useful, but in case the pupil encounters an unknown task, he will restrict only on finding the analogy, he won't avoid mistakes, and usually he will not find the solution. Trying to find unknown solutions pupils have to be accomplished and advised, absolutely understand the essence of the task and the essence of solution. We have to admit that teachers pay too little attention for comprehensive analysis of the task, for organizing the search of the task solution, especially – for the conclusive analysis, where we find out what conclusions we can draw out of the solution.

References

[1] B. Balčytis, *Skaičių šalis*, IV klasė, Šviesa, Kaunas (1999).
[2] B. Bitinas, *Statistiniai metodai pedagogikoje ir psichologijoje*, Kaunas (1974).
[3] *Bendrojo išsilavinimo standartai*, I–X klasės, Vilnius (1998).
[4] A. Bühl, P. Zöfel, *Professionelle Datenanalyse mit SPSS für Windows*, Bonn, Paris (1996).
[5] R. Harvey, A. Hammer, Item response theory, *Counseling Psychologist*, **27**(3), 353–383 (1999).
[6] A. Kiseliovas, D. Kiseliova, *Matematikos pasaulyje*, IV klasė, I, II knygos, Alma Littera, Vilnius (2000).
[7] A. Kiseliovas, D. Kiseliova, *Matematikos teminiai testai*, Alma Littera, Vilnius (2002).
[8] *Lietuvos bendrojo lavinimo mokyklos bendrosios programos*, Leidybos centras, Vilnius (1994).
[9] R. McCall, *Fundamental Statistics for Behavioral Sciences*, Harcourt Brace & Company (1994).
[10] *Standards for educational & psychological tests*, Washington (1999).
[11] А. Анастази, С. Урбина, *Психологическое тестирование*, 7-е международное издание, Санкт-Петербург, Москва-Харьков-Минск (2001).
[12] Б. Битинас, *Многомерный анализ в педагогике и педагогической психологии*, Вильнюс (1971).
[13] Б. Битинас, Р. Паулавичюс, *Система анализа педагогических данных на ЭВМ*, Вильнюс (1987).

Matematinių gebėjimų taikymų testo užduočių validacija

D. Kiseliova, A. Kiseliovas, I. Donielienė, O. Kiseliova

Straipsnyje analizuojamos ketvirtų klasių moksleivių baigiamojo matematikos taikymų testo užduočių diagnostinės kokybės charakteristikos: validumas, sunkumas, x^2, Gutmano patikimumo koeficientas ir t. t. Testo užduočių vidinė struktūra papildomai nagrinėta alfa faktorinės analizės metodu. Tyrimo imtis – 361 ketvirtokas iš įvairių Lietuvos miestų ir rajonų mokyklų.