Edukacja
An interdisciplinary approach 2013, 1
ISSN 0239-6858
pp. 59–69

Double-digit coding of examination math problems

Agnieszka Sułowska, Marcin Karpiński
Mathematics Section, Educational Research Institute*

Various methods are used worldwide to evaluate student solutions to examination tasks. Usually the results simply provide information about student competency and after aggregation, are also used as a tool of making comparisons between schools. In particular, the standard evaluation methods do not allow conclusions to be drawn about possible improvements of teaching methods. There are however, task assessment methods which not only allow description of student achievement, but also possible causes of failure. One such method, which can be applied to extended response tasks, is double-digit coding which has been used in some international educational research. This paper presents the first Polish experiences of applying this method to examination tasks in mathematics, using a special coding key to carry out the evaluation. Lessons learned during the coding key construction and its application in the assessment process are described.

Keywords: evaluation, double-digit coding, final exam, PISA, TIMSS.

Insufficient use of information from national examinations

In most European countries, national examinations are mainly used to assess student achievement at the end of a certain stage of education, as well as to monitor and evaluate schools or even the whole education system. Some countries (e.g. France, Denmark, Sweden, Hungary) claim that their national examinations are aimed at identification of educational needs and identification of appropriate teaching methods (Eurydice, 2009). It could be inferred from those claims that formative assessment is one aspect of those exams.

The problem presented in this article arises from the observation that national examinations in most countries, in their present form, usually do not serve this purpose. We agree with the view that their main role is really to summarise the achievements of students

The article was written on the basis of the study Diagnoza kompetencji gimnazjalistów 2011 (Diagnosis of the competences of lower secondary school students 2011) carried out within the systemic project „Quality and effectiveness of education – strengthening of institutional research capabilities" executed by the Educational Research Institute and co-financed from the European Social Fund (Human Capital Operational Programme 2007–2013, Priority III High quality of the education system). This article was published primarily in Polish language in *Edukacja*, *119*(3) 2012.

* Mail address: Agnieszka Sułowska, Pracownia Matematyki, Instytut Badań Edukacyjnych, ul. Górczewska 8, 01-180 Warszawa, Poland. E-mail: a.sulowska@ibe.edu.pl

at the end of a stage of education. There is nevertheless, the possibility that they may also perform a formative assessment function. This possibility is illustrated with an empirical example. Exams prepared in this way can contribute to the more effective development of students' skills and the gradual improvement of the quality of teaching methods.

However, to interpret the exam results to make assessment and improvement of teaching methods possible, an appropriate analysis of student solutions to exam tasks is needed. The outcomes of this analysis need to be coded for statistical tools to be used. Tools and methods designed for such an analysis may also be used by teachers themselves to examine their students' reasoning and to identify causes of their failures in order to improve their own teaching methods. In the long run they may be able to raise standards.

A review of final mathematics examinations in various countries revealed that, just like in Poland, neither the methods of coding solutions, nor the means of communicating exam results allowed a deeper analysis of causes of student failure or improvement instruction methods. Here are some typical examples.

France. *Diplôme nationale du brevet* is a national examination, but despite its central administration and marking, assessment, and interpretation of results are not uniform in practice (Eurydice, 2009). A mathematics exam paper (*série collège*) contains around 20 tasks, most of which are extended--response tasks which are marked by assigning points for stages of the solution. A separate 4 points are assigned for layout and method of presentation in the whole work (the maximum total number of points for the tasks is 36). Exam results serve mainly to monitor schools and are presented in a report on students' basic mathematical competencies.

The Netherlands. The VWO exam corresponds to the Polish upper secondary school leaving exam (*matura*), but as the criteria for selection of students at an earlier stage of education are more demanding than in Poland, only 20% of Dutch students reach this point. The mathematics paper only contains extended-response tasks (over 20 items). In marking tasks points are assigned to a sequence of rather meticulously identified stages of solution. In the marking scheme, these stages are described in detail, usually in terms of the results of partial calculations. The exam results are mainly used for assessment of student achievement.

Russia. A state exam after grade 11. The exam paper comprises 20 tasks divided into two groups. For each of 14 tasks from the first group, the student may receive at most 1 point. Although the tasks have the form of extended-response tasks, the student is not required to present complete reasoning. The final answer, which is always an integer or a decimal fraction, is sufficient. For the remaining six tasks, the student has to give the full solution. They are marked by awarding points to each stage of the solution. Description of those stages is so general that it covers various approaches to the solution.

The systems of the International Baccalaureate (IB) and the International General Certificate of Secondary Education (ICGSE) examinations. Similar task coding methods are used in both these systems. For each task, the examiner assigns codes comprised of letters and digits. The letter defines the type of skill demonstrated by the student and the digit indicates the level achieved. The marking scheme describes which letters and digits should be used to complete each task. Letter designations include: M – use of the right method, A – providing the right answer, R – presentation of the right reasoning, G – obtained by means of a graphic calculator. In the IB examination, each examiner, besides coding of exam papers, is also asked to provide a report containing representative examples of student solutions (correct and incorrect). The examples are used to develop parts of the official exam report entitled "Recommendations and guidelines for teaching future exam takers". The guidelines are, however, extremely brief and aimed at obtaining a better exam result, rather than improvement of teaching methods.

Proposal of a solution: double-digit coding

Traditionally, in school practice, in external examinations and in educational research, while testing the mathematical knowledge and skills, student solutions are checked by means of criterion-referenced assessment. In general, a score is assigned to each solution in accordance with criteria described in the marking scheme.

Double-digit coding is a variant of criterion-referenced assessment. It combines traditional marking with collecting additional information. Double-digit coding assigns two scores to each solution. The first score, just as in traditional marking, corresponds to the assigned number of points or general level of correctness of the solution. The second digit indicates the method used by the student to solve the task, the reasoning or strategy used by the student or the kind of error made. Application of this method of assessment requires preparation of a marking scheme called a coding key – a system of available codes and their exact descriptions, followed by training of examiners or coders (Dossey, Jones and Martin, 2002).

Double-digit coding has been employed in international surveys TIMSS and PISA (OECD, 2005; Olson, Martin and Mullis, 2008). In each of those cyclic international research programmes, some extended-response tasks were marked using of double-digit coding. International reports from subsequent editions of those surveys contain numerous analyses and much information concerning all participating countries. However, there is no analysis of the results from double-digit coding, probably due to the high level of generality at which the analyses are performed and communicated (Mullis, Martin and Foy, 2008; OECD, 2004).

Analysis of detailed data is performed on methods used to solve specific tasks, reasoning, strategies and types of errors made by students i.e. the data provided by double-digit coding. This makes more sense at a level of lower generality, e.g. in comparative analyses carried out for a few countries, national surveys or analyses at school or class level.

Examples of analyses performed with the use of double-digit coding

An example of such a comparative analysis carried out for a group of Nordic states is presented in *Northern lights on PISA. Unity and diversity in the Nordic countries in PISA 2000* (Turmo, Kjærnsli and Pettersson, 2003). The authors present a task entitled "Antarctica" used in the PISA 2000 survey and offer a detailed discussion of its coding. In the task, specific codes indicate the solution method and any error made. The authors compare the results obtained for that task in each of five Nordic countries (Denmark, Finland, Iceland, Norway, Sweden) both at a general level and at the code level. They analyse the frequency specific methods to solve the task in each country. They also offer hypotheses to explain why the task was found to be so difficult in all countries and why so many students avoided the task.

The publication mentioned above is based on the PISA 2000 survey. Unfortunately, an analogous publication concerning later editions of the survey could not be found, although it had been promised. Norwegian and Finnish reports on PISA 2003, PISA 2006 and PISA 2009 surveys or their shorter versions available in English do not contain any reference or analysis based on the results of double-digit coding.

An example of the use of information available from the double-digit coding introduced internationally is offered by *Learning mathematics for life. A perspective from PISA* (OECD, 2010b). In the chapter "Mathematical problem solving and differences in students' understanding", just like in the publication quoted above, the authors present one task, "Steps", used in the PISA 2003 survey and provide a detailed discussion of its coding. In that task, specific codes indicate the types of errors made. The authors analyse the frequency of specific errors at the level of the whole survey, as well as in specific countries. For instance, they note that Poland belongs to the few countries, where students make errors in conversion of units more often than in other OECD countries. An analysis of the frequency of specific codes in selected countries is also used by the authors to verify hypotheses concerning similarities between countries in the scope of mathematical skills.

In the chapter summary, the authors emphasise the difficulty of analysing solution methods, strategies or errors with regard to specific tasks on the basis of the PISA survey data, due to the very limited number of tasks coded with two digits. They stress that PISA is a mass survey carried out on such a large scale that makes it impossible to gather all the necessary data on solution methods and reasoning. They also encourage the use of such tasks in every-day school work, where they may be additionally developed into a discussion and students may be asked to provide arguments. The authors also stress that deeper and more detailed analyses may be carried out at a national level and their results compared with those of international surveys.

In Polish surveys concerning the results of the lower secondary school leaving exam, attempts at analyses similar to double-digit coding can also be found. In 2005, in the Regional Examination Commission in Cracow, an open task coding system was prepared, in which, besides points for solving a task, a separate letter code was allocated to various categories of errors made by students (Kołodziej, 2007).

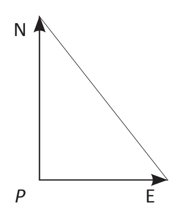
The use of double-digit coding in the study *Diagnoza kompetencji gimnazjalistów 2011*

Here we present the application of the double-digit coding in the study *Diagnoza kompetencji gimnazjalistów 2011* (Diagnosis of the competencies of lower secondary school students 2011 – DKG): one of the extended-response tasks used in the study, various ways of solving it and, most importantly, the coding key used for marking it. We also present the experience gained during the coding key construction and in the process of its application and the lessons learned.

In the general study DKG, performed by the Central Examination Board (*Centralna Komisja Egzaminacyjna* – CKE) in 2011, students from almost all lower secondary schools (approx. 7000) in Poland participated. The problems were prepared by the CKE. Students' work was marked at schools by teachers using the traditional marking scheme prepared by the CKE. In the research part of the DKG, implemented by the Education Research Institute (*Instytut Badań Edukacyjnych* – IBE), students solved the same sets of tasks as in the general DKG. However, students' work was marked and coded by trained external examiners according to marking schemes (coding keys) prepared by the IBE. One of the three extended-response tasks used in the survey is shown below.

Task 21. "Fishing boats" (the first extended-response task in the set)

Two fishing boats depart from the harbour (point *P*) at the same time: one goes North with the constant speed of 4 knots, the other goes West with the constant speed of 3 knots.



Calculate the distance between the boats two hours after departure. Provide the result in kilometres. Write down all calculations. To solve the problem use the following information: 1 knot equals 1 sea mile per hour, 1 sea mile = 1852 m.

All three extended-response tasks in the set assessed mastery of the so-called complex skills. They include the ability to use or create a problem-solving strategy and the ability to reason and present arguments. Task 21 assessed mastery of the ability to create an appropriate solution strategy. The importance of complex skills in education is particularly emphasised by the new Polish core curriculum. It poses a new challenge to creators of tasks and marking schemes (coding keys), to prepare them in a way that exposes, captures and appropriately assesses the reasoning or strategies presented in the solution.

Task solution methods

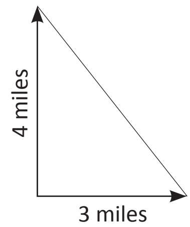
Solution of the task is in three steps:

1. calculation of the distance between fishing boats applying the Pythagorean theorem;
2. multiplication of the result by 2, to account for the travel time of 2 hours;
3. conversion of miles into kilometres.

These steps may be carried out in any order. In the first solution (I) presented below, the steps are made in the above order: 1, 2, 3. In the second solution, first step 2 was performed, then step 1 and finally step 3, whereas in the third solution, the first is step 3, then step 2 and finally step 1. The use of the right strategy to solve the problem consists of the selection of the right order of performing the steps, although every order leads to a correct solution, the difficulty of the task depends significantly on the choice made.

1st solution method:
One hour after departing: the first fishing boat covered 4 sea miles, the other boat covered 3 sea miles. The directions in which the boats travelled were perpendicular, so the Pythagorean theorem can be used. Using the specific case of the Pythagorean theorem, the so called Egyptian triangle (a right-angled triangle with the sides of 3, 4, 5) we know that the distance between the boats one hour after departure is 5 sea miles.

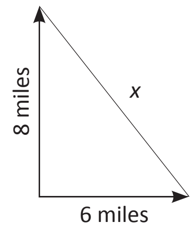
Two hours after departure, the distance between the boats will be twice as big, that is it will be $2 \cdot 5 = 10$ sea miles. 10 sea miles equals $10 \cdot 1852$ m = 18 520 m = = 18.52 km.



This is the optimum strategy to solve the task – basically it does not require any calculations. Using this strategy, the problem can be solved mentally.

2nd solution method:
Within two hours, one of the fishing boats covered $2 \cdot 4 = 8$ sea miles, the other $2 \cdot 3 = 6$ sea miles. The distance between the boats (x) is calculated by applying the Pythagorean theorem.

$x^2 = 8^2 + 6^2$
$x^2 = 64 + 36$
$x^2 = 100$
$x = 10$ (sea miles)
10 sea miles equal
$10 \cdot 1852$ m = 18 520 m =
= 18.52 km



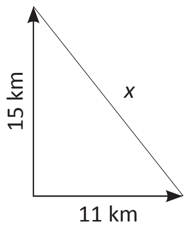
This method of solving the problem is also very good, although arriving at the distance between the boats requires some calculations. The simplicity and efficiency of this solution results from the fact that – just like in the prior solution – distances expressed in miles, that is one-digit integers, are plugged into the Pythagorean theorem.

3rd solution method:
1 knot is 1 sea mile per hour, that is 1.852 km/h. Thus, the first fishing boat will cover 7.408 km, and the other 5.556 km during one hour. After two hours, it will be 14.816 km (approx. 15 km) and

11.112 km (approx. 11 km). Applying the Pythagorean theorem, we calculate the distance between the boats after 2 hours:

$x^2 = 15^2 + 11^2$
$x^2 = 225 + 121$
$x^2 = 346$
$x =$
$18 < x < 19$

The distance between the fishing boats is around 18.5 kilometres.



This is the worst and, at the same time, the most routine solution strategy. The routine manifests itself here in the mindless application of the rule: first convert all provided distances into kilometres. As a consequence, two five-digit numbers are obtained, which must be squared when applying the Pythagorean theorem. However, students could not use calculators in the exam or during the survey, so it was practically impossible. The only feasible way to avoid this was to round the obtained numbers, for example to full integers. However the use of approximations is neither easy nor obvious for students and they often have doubts whether it is appropriate to do so.

The next difficulty was encountered by students when applying the Pythagorean theorem. They obtained 346 which is not a square of a natural number and were supposed to find the square root of it. In this situation, some of the students decided to provide the distance in the form of “km”, which is not a particularly useful piece of information and it contradicts the practical context of the problem. The only sensible way of providing the distance between the fishing boats was, therefore, estimation, which also presented difficulties at lower secondary school level.

Assumptions and limitations of the coding key

All coding keys prepared for use in this study had been developed to make possible the extraction and collection of data about the problem-solving method, the reasoning or strategy and any errors made. However, the result of the application of the developed keys in terms of student marking had to be exactly the same as the result obtained from the marking schemes published by the CKE. A student, who received *x* points for his or her solution according to the CKE marking scheme had to score exactly the same after application of the coding key used in the survey.

CKE’s marking scheme in terms of degree of task solution

Performance level:

P_6 – “complete solution” – 3 points: calculation of the distance between the fishing boats two hours after departure in km (18.52 km);

P_4 – “the inherent difficulty of the problem was overcome without errors, but the solution is incomplete or a part of the solution contains substantial errors” – 2 points: calculation of the distance between the fishing boats two hours after departure in miles (10 sea miles);

P_2 – “significant progress is achieved but the inherent difficulties of the problem were not overcome” – 1 point: calculation of the distance travelled by each fishing boat two hours after departure (8 sea miles, 6 sea miles), or calculation of the distance between the fishing boats one hour after departure (5 sea miles);

P_0 – “solution that does not constitute progress” – 0 points: incorrect solution or no solution (CKE, 2011).

Coding key developed for the survey

Category 3. Complete solution: calculation of the distance between the fishing boats two hours after departure in km – 3 points.

- code 3.1 – 18.52 km – values in miles put into the Pythagorean theorem (6 and 8 miles or 3 and 4 miles);

- code 3.2 – the number in decimal form with value between 18 km and 19 km, such as “around 18.5 km” – approximated values in kilometres were put into the Pythagorean theorem (e.g. 15 km and 11 km or 14.8 km and 11.1 km);
- code 3.3 – a root or expression containing a root, with value between 18 km and 19 km, e.g. $\sqrt{346}$ km or $2\sqrt{86}$ km – approximated values plugged into the Pythagorean theorem in kilometres (e.g. 15 km and 11 km or 14.8 km and 11.1 km);
- code 3.4 – $\sqrt{342.9904}$ km or $2\sqrt{85.7476}$ km – values plugged into the Pythagorean theorem in kilometres without approximation (14.816 km and 11.112 km or 5.556 km and 7.408 km).

Category 2. The inherent difficulties of the problem were overcome, but the solution was not finished or the later part of the solution contained errors – 2 points.

- code 2.1 – calculation of the distance between the fishing boats two hours after departure in miles (10 sea miles);
- code 2.2 – calculation of the distance between the fishing boats one hour after departure in metres or kilometres (9260 m or 9.26 km);
- code 2.3 – solution of the problem up to the end (calculation of the distance between the fishing boats two hours after departure) but with an arithmetic error or with an error in conversion of units.

Category 1. Significant progress was made but the inherent difficulties of the problem were not overcome – 1 point.

- code 1.1 – calculation of the distance covered by each of the fishing boats within two hours in miles (8 sea miles, 6 sea miles);
- code 1.2 – calculation of the distance covered by each fishing boat within two hours in metres (14,816 m, 11,112 m) or kilometres (14.816 km and 11.112 km);
- code 1.3 – calculation of distance between the fishing boats one hour after departure in miles (5 sea miles).

Category 0. Solution does not constitute progress – 0 points.

- code 0 – incorrect solution;
- code 9 – no solution.

Application of the coding key

When using the coding key that had been prepared for marking students’ solutions in the survey, a type of solution emerged for which no appropriate code existed. Also, two types of solutions that were significantly distinct from the perspective of students’ skills shown were not distinguished by the coding key. The classification of incomplete solutions (like those described under codes 2.1 or 2.2), with an additional arithmetic error or an error in converting units from metres into kilometres was evidently problematic. The general rule adopted when developing the coding key was that a solution with an arithmetical error should be classified into the category below the category describing the correct analogous solution. According to this rule, an incomplete solution with an arithmetical error should drop from category 2 to category 1 and there should be an appropriate code for it in that category.

Another problem that emerged was excessive “capacity” of code 1.2. It covered situations in which students either did not know how to solve the task and performed only the routine supporting calculations, or solutions in which there was full, correct reasoning which were not completed due to the wrong strategy and the consequent arithmetical difficulties. Although the end effect in both situations was the same, the student could not calculate the distance between the fishing boats, both types of attempted solution should be distinguishable in terms of the skills demonstrated by the student.

Analysis of coding

Table 1 presents percentages of solutions classified under specific categories and specific

Table 1
Shares of types of solutions of the "Fishing boats" task broken down into codes

Code	3.1	3.2	3.3	3.4	2.1	2.2	2.3	1.1	1.2	1.3	0	9
[%]	9.9	0.7	2.1	0.4	0.6	3.5	9.3	4.1	23	1.2	41.2	4

codes: from category 3 – full solutions, to category 0 – incorrect solutions. Code 9 indicates student's failure to attempt the task.

Category 3
Only 13% of students fully solved the task, that is obtained a code from category 3 for their solution. Almost 10% students received code 3.1, i.e. they solved the task in the optimum, non-routine way, substituting distances expressed in miles covered by the fishing boats into the Pythagorean theorem. Those were the students who could think unconventionally – they did not fall into the trap of following the traditional but this time non-optimal path. Some of them found their own way to solve the task from the outset. Others started in the traditional way from conversion of miles into kilometres, but having seen where the path led were able to abandon it and start solving it again, looking for a new, better way.

One third of the students (3.2%) solved the task using a more obvious, but worse strategy. These students first converted the distances covered by each of the fishing boats from miles into kilometres, and only then obtained values (precise or approximate) to plug into the Pythagorean theorem. Considerably fewer students embarked on this solution, since, as already mentioned, this is the traditional path. However, only 3.2% of all students solving the task managed to reach the goal following this route. Of those students, only 4 out of 1000 received code 3.4, i.e. solved the task correctly, putting values without approximations into the Pythagorean theorem, i.e. raising to a power and extracting a root of five-digit numbers without a calculator. They were students so skilled in arithmetic that they could even complete such a complicated task successfully. It did not testify to their critical thinking skills but rather to their lack of thinking combined with arithmetic proficiency.

More students, 21 out of 1000, received code 3.3. They approximated the obtained values before substituting them into the Pythagorean theorem but they could not or did not see the need to estimate the answer and provided it in the form of a root. This procedure is, in our opinion, incorrect in the task embedded in such a practical context, although it does not constitute a flaw in reasoning. Unfortunately, most of those 3.2% students who used the worse solving strategy successfully, did just that. Very few students, 7 out of 1000, received code 3.2. They approximated values before substitution into the Pythagorean theorem and estimation of the value obtained in a practical form, such as "around 18.5 km". In our opinion, of the 3.2% of students, the 7 students managed best – they realised that the approximation which they provided was needed.

Category 2
In category 2 codes 2.1 and 2.2 cover the solutions of students who overcame the inherent difficulty of the task, i.e. correctly used the Pythagorean theorem to calculate the distance between the fishing boats, but did not complete the calculations, a total of 4.1%. Six out of 1000 only lacked the conversion of the calculated distance between the boats from miles into kilometres. Most of them (35 out of 1000), calculated the distance but one hour after departure, rather than two hours. What was missing, therefore, was only multiplying the value by 2. These students were either

absent-minded and forgot about this simple operation, or, more likely, did not fully understand the situation presented in the task and as a result, did not realise that they should multiply the result by 2.

In the same category, code 2.3 covers the results of students who completed the task but made arithmetic errors or errors in conversion of units. Such students comprised 9.3% of the sample. This suggested that almost every tenth student knew how to solve the task but did not obtain the right result and as a consequence, did not obtain the full score because of an arithmetic error. As mentioned above, most of the errors resulted from substituting kilometres into the Pythagorean theorem without approximation. An arithmetical error in such a case is not unusual. This is simply the consequence of choosing a bad solution strategy.

Category 1

The total of 28.3% students were in this category. They started solving the task by undertaking some reasonable steps but they were not able to overcome the inherent challenge of the task, i.e. to apply the Pythagorean theorem correctly. The students, whose solutions were classified with codes 1.1 and 1.3, performed only one small step on the path to solving the problem. Under code 1.1, it was to multiply the two values by 2. There were 4.1% of such students. Under code 1.3, it was to convert of miles into metres or kilometres. There were 1.2% of such students.

The great majority of students (23%), whose solutions were in category 1 received code 1.2. These students correctly performed the two supporting steps of the solution, i.e. they converted the values provided into metres or kilometres and multiplied them by 2, i.e. they calculated the distance covered by each of the fishing boats in 2 hours in kilometres. Further calculations were missing, incorrect or started but not completed. A large proportion of students who received that code subtracted the calculated distances, which was, of course, wrong. Other students, who also received code 1.2, substituted the obtained distances in kilometres without using approximations into the Pythagorean theorem. Unfortunately, raising five-digit numbers to the second power without a calculator exceeded their abilities and they abandoned the calculations. This means that these students perceived the path to solve the task but were unable to follow it, as, although correct, it was arithmetically too difficult for them.

The result in both situations was, unfortunately, the same. Students were not able to calculate the distance between the fishing boats. As a consequence, both types of solutions fell into category 1, which meant that their authors scored only 1 point. The second type of solution described above, is a particularly dramatic example of adopting a bad strategy.

Category 0

This category contains only two codes: 0 and 9. Code 0 was given to students who only converted the provided values into metres or kilometres and stopped at that, as well as students whose solutions were simply incorrect. In both cases these solutions did not demonstrate sufficient progress towards a solution. Unfortunately, as many as 41.2% of all students solving the task presented such solutions. Code 9 was given to students who did not even attempt the task. This was only 4% of students. That suggests that almost all students perceived the task as within their abilities.

The above analysis of the results of coding confirms that the choice of the right strategy was essential to solve the task successfully. In consequence, a good strategy was chosen by 11.7% of students. This is not a poor result, considering that it required the student to independently diverge from a routine that could seem initially to be the right one. The vast majority of those students solved the task correctly unlike students following the routine path which usually led to them giving incorrect answers. Discussing such an example with students may give them the stimulus to take a new look at strategy for solving other problems in future.

Conclusions

All three tasks included in the DKG set tested the mastery of complex skills. These included the ability to use or create a problem-solving strategy and to reason and present arguments. The task "Fishing boats" described above verified mastery of the ability to develop an appropriate solving strategy. The importance of complex skills in education is particularly emphasised by the new Polish core curriculum. It poses a new challenge for the creators of tasks and marking schemes (coding keys), which is to prepare them in a way that brings out, captures and appropriately assesses the reasoning or strategies presented in the solution.

The present most common method of marking examination tasks provides a single message to students and teachers: the level of performance. As a result, the impact of examination results on modification of the teaching is limited to increasing the number of exercises of the kind which caused difficulties at the exam. This leads to the phenomenon of studying for the exam. The double-digit marking scheme proposed provides an opportunity to provide much richer information to the education system. It enables recognition of the most important kinds (and sometimes causes) of errors made by students and allows the identification of areas of student knowledge where they resort to irrational approaches. This method of coding also allows testing of solution strategies chosen by students and the quality of reasoning. It provides a basis for better teaching methods to help students avoid errors.

Literature

CKE (Centralna Komisja Egzaminacyjna – Central Examination Board) (2011). *Badanie diagnostyczne w roku szkolnym 2011/2012. Część matematyczno -przyrodnicza. Matematyka. Odpowiedzi i propozycje oceniania zadań* [*Diagnostic test in the school year 2011/2012. Mathematics and science. Maths. Answers and suggestions of task assessment*]. Retrieved from http://www.cke.edu.pl/images/stories/000011_gim_pr/matem/GM-M1.zip

Dossey, J. A., Jones, C. O. i Martin, T. S. (2002). Analyzing student responses in mathematics using two-digit rubrics. In D. F. Robitaille, A. E. Beaton (eds.), *Secondary analysis of the TIMSS data*. Dordrecht: Kluwer Academic Publishers.

Eurydice (2009). *National testing of pupils in Europe: objectives, organisation and use of results*. Bruxelles: The Education, Audiovisual and Culture Executive Agency (EACEA) – Eurydice.

Foy P. and Olson J. F. (eds.). (2010). *TIMSS 2007 user guide for the international database*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

Kołodziej, K. (2007). Kategoryzacja rozwiązań zadań otwartych części matematyczno-przyrodniczej egzaminu gimnazjalnego. Materiały z XIII Konferencji Diagnostyki Edukacyjnej [Categorisation solutions of open tasks in mathematics and science of lower secondary school exam. Proceedings of the XIII Conference on Educational Diagnostics], Łomża.

Mullis, I. V. S., Martin, M. O. and Foy, P. (2008). *TIMSS 2007 international mathematics report: findings from IEA's trends in international mathematics and science study at the fourth and eighth grades*, Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

OECD [Organisation for Economic Co-operation and Development] (2004). *Learning for tomorrow's world – first results from PISA 2003.* Paris: OECD.

OECD (2005). *PISA 2003. Technical report,* Paris: OECD.

OECD (2007). *PISA 2006 science competencies for tomorrow's world (volume I).* Paris: OECD.

OECD (2009). *PISA 2006. Technical report.* Paris: OECD.

OECD (2010a). *PISA 2009 results: what students know and can do. Student performance in reading, mathematics and science (volume I).* Paris: OECD.

OECD (2010b). *Learning mathematics for life. A perspective from PISA*. Paris: OECD.

OECD (2012). *PISA 2009, Technical report.* Paris: OECD.

Ruddock, G. J., O'Sullivan, Ch. Y., Arora, A. and Erberber, E. (2008). Developing the TIMSS 2007 mathematics and science assessments and scoring guides. In J. F. Olson, M. O. Martin, I. V. S. Mullis, (eds.), *TIMSS 2007 technical report* (pp. 30–32), Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Turmo A., Kjærnsli M. and Pettersson A. (2003). Mathematical literacy and competency classes. In S. Lie, P. Linnakylä, A. Roe (eds.), *Northern lights on PISA. Unity and diversity in the Nordic countries in PISA 2000.* Oslo: University of Oslo.