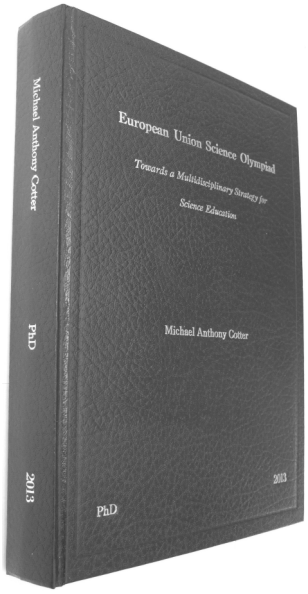
Olympiads in Informatics, 2013, Vol. 7, 186–190

REVIEWS, COMMENTS



European Union Science Olympiad

Towards a Multidisciplinary Strategy for Science Education

Author: Michael Anthony COTTER
Publishing house: Dublin City University
Country: Ireland
Year of edition: 2013
Language: English
Number of pages: 303
ISBN: N/A

This PhD thesis, by Michael A. Cotter, is primarily practice-focused research on a progressive version of science education with historical, social and political dimensions. It tracks the development and organisation of a unique science competition for sixteen year old students, the European Union Science Olympiad (EUSO) from its conception to the end of its tenth year. The motivation for establishing the EUSO was to develop, organise and maintain over an extended period of time a science competition which would develop "Team Science Tasks" that would integrate science, be problem-based, be connected to the real world and involve the construction of knowledge, higher-order thinking, alternative solutions, depth of knowledge and sophisticated communication between the team members. The EUSO was designed to fill a gap in the science olympiad landscape, whilst simultaneously functioning as a stimulus for increased interest in and enjoyment of scientific phenomena at a critical period in students' education.

A crucial element in this success story has been the methodology used in the organisation of the EUSO – *Participatory Action Research* (*PAR*). Through this democratic process, participants were released from the constraints of the established olympiad structures and enabled to function in an environment which allowed them to investigate their

own reality in order to change it. Such actions have resulted in the development of a new science olympiad model, the EUSO model.

The idea of the EUSO crystallized for the researcher following his attendance at and his experience of organising Ireland's participation in the 6th International Olympiad in Informatics (IOI), in Haninge, Sweden in 1994. At this event, he met and spoke with the then President, Dr. Yngve Lindberg (1900–1966) and with the Chairman of the Scientific Committee for IOI 1994, Dr. Håkan Strömberg. This enabled him to organise Ireland's participation in the IBO, IChO and IPhO in subsequent years.

In the early years, Irish students' success stories emanated from the IOI, primarily through the establishment of a junior section in the Irish Informatics Olympiad and to the setting up of online computer programming clubs and training programmes by Mr. Charlie Daly, the Irish team leader. As a result, Eóin Curran was awarded a silver medal at IOI 1997 and Martin Orr a gold medal at IOI 2003. These results suggested that Ireland's performance at other olympiads could be enhanced through early intervention.

Much of the documentation emanating from the EU and elsewhere signalled a consensus regarding the features of science education that EU member countries had in common. These included:

- decline in the numbers studying science in secondary schools;
- sharp decline in the numbers studying chemistry and physics;
- steep decline in the number of girls studying chemistry and physics;
- lack of interest among students in school science;
- poor attitude towards science and scientists;
- shortage of suitably qualified science teachers, particularly in chemistry and physics;
- ineffectual teaching of science;
- single subject science curricula;
- lack of integration in school science;
- lack of experiments / practicals in school science;
- shortage of university science graduates.

The researcher was fully aware that his personal perspectives and preconceptions would have a bearing on the research process and his approach to answering the research question. In this regard he realised that he subscribes to the view that the kind of knowledge that is valid and satisfactory is that which is created through the subject's interactions with the world. However, subjects construct their own meaning in different ways and experience the world from different perspectives. Such a constructivist stance calls for an interpretivist approach to knowledge creation. In interpretivism, the *'world is interpreted through the classification schemas of the mind'* (Gray, 2006, p. 20); the emphasis is on understanding the real workings behind 'reality'. For the researcher, attempts to understand this reality were grounded in people's experiences of that reality. The task was to explore people's multiple perspectives in the natural settings.

The researcher was also mindful of the fact that his chosen methodology should align with his epistemological stance and theoretical perspective, i.e., constructivism and interpretivism, respectively. In this regard, the postmodernist approach to knowledge cre-

ation promoted by Action Research (AR) was considered appropriate. AR enables the researcher to:

> *'Develop a context in which individuals and groups with divergent perceptions and interpretations can formulate a construction of their situation that makes sense to them all – a joint construction'* (Stringer, 1999, p. 45).

For the researcher, the particular AR model chosen, namely Kemmis and McTaggart's (2003) Participatory Action Research (PAR) was selected due to its primary focus: authentic participation. In addition, the emphasis on investigation of actual practices and the concentration on transformation of practitioners' practices in an egalitarian manner enabled adoption of PAR as the model most likely to have the capacity to handle the variety of challenges that the EUSO concept would inevitably generate.

However PAR does not cover all aspects of this research. The historical aspect of the EUSO is a crucial aspect and a vital component of the narrative.

The researcher's direct involvement in the IBO, IChO, IPhO and the IOI gave him a unique insight into their structures, organisation and management and gave him access to a wide range of influential people who gave him their views on aspects of the events and the profile a new olympiad might adopt. This raised ethical issues around privileged information and how it might be used. However the researcher was open and not clandestine or covert in any way and as many opinions as possible were sought. Cooperation and collaboration were ensured and dissenting opinions were listened to. Gray (2006) citing Badger (2000) suggests that

> *"At least superficially Action Research seems to pose few ethical dilemmas because it is based on a philosophy of collaboration for the mutual benefit of the researchers and participants"* (p. 388).

While Action research is not a "smash and grab" approach to research or what Lather (1986) calls "rape research" (p. 261), it requires negotiated access, confidentially and the right to withdraw.

This thesis reviews the limited literature on the history of the international science olympiads from the Leningrad Mathematical Olympiad (LMO) in 1934 to the present day olympiads by using primary sources where possible. Literature on relevant topics raised by the research question such as, why teach science are explored. The central role of science education in the education systems across the EU is investigated, in particular the perceptible decline in participation in the Senior Cycle of the secondary education systems in chemistry and physics by all students, but especially girls. This has raised concerns about the future supply of well qualified scientists, engineers and science teachers. This has led to much research on the pivotal role of the teacher, their qualifications and training, the curriculum content of the subjects they teach and the teaching methods that are employed and how this in turn influences the interest and attitude of students. The critical role of assessment is reviewed.

Chapter 4 of this thesis outlines the research methodology chosen. The Epistemological Stance and Theoretical Perspective of the researcher are presented as is the rationale

for choosing Participatory Action Research (PAR), because of its aim to transform and to be participatory, practical, collaborative, emancipatory and reflexive. While acknowledging the overarching importance of the PAR approach, the import of the setting and historical account of the EUSO over a ten year period is emphasised. The advancement of the Action Research methodology from the early work of Kurt Lewin (1890–1947) through its many modifications and developments to PAR developed by Kemmis and McTaggart (2003) is described. The challenges faced and overcome, the data collection method of informal and formal meetings, diaries and interviews are graphically presented in five cycles over a five year period. The cyclical nature of the research is further described in the development of the EUSO constitution from 2002–2012

The thesis presents a brief summary of the twenty EUSO Tasks developed over a ten year period 2003–2012. It also serves as a historical record which describes how these unique Team Science Tasks, a central feature of the EUSO were created. As well as each summary it also illustrates how each Task contributes, in its own way, to the concept of "Rich Tasks" in a progressive version of science education. Each task was reviewed to see if it conforms to Progressive Pedagogies (Hayes *et al.*, 2006). In line with PAR, it describes how the Scientific Committee from the host country chooses the topics and designs the Task before presentation to the mentors at a General Assembly (GA) meeting. All aspects are discussed before finalisation by task designers and mentors in a collaborative manner. The need for future research into the amount of interaction that takes place between team members is alluded to. The elimination of high-stakes assessment is emphasised as a feature of the EUSO in the presentation of the results. This feature developed during the PAR process, by getting the mentors to emancipate themselves from the familiar concept of olympiads where the emphasis is on winning and not on taking part.

The results from the EUSO are presented in two five-year periods because of EU enlargement in 2004 and 2007. This contributes to the historical record of the EUSO and highlights some of the trends that have developed. It was expected that the former Soviet Bloc countries, because of their success at the international olympiads would perform well. The interviewing, which was one of the research methods used in PAR was continued to help the researcher comment on the results. In the first five years the former Soviet Block countries featured strongly in the gold and silver medal categories while the early EU members performed well in the silver category and were dominant in the bronze category. In the years 2008–2012 the former Soviet Bloc countries were dominant in the gold and silver categories while the early EU members featured strongly in the bronze category.

The final chapter, Conclusions & Recommendations, points to the success of the EUSO as the product of Participatory Action Research (PAR). Ten years of the EUSO has seen the number of participating countries and the numbers of students steadily increase. Ten EU governments have hosted the EUSO between 2003–2012 and the science faculties of the universities in these countries have cooperated to develop twenty, integrated, content rich, team science tasks. Finally recommendation including the promotion of an integrated science curriculum and the possibility advanced programmes in science. Areas where additional areas of research might be undertaken are highlighted.

References

Cotter, M., Petersen, S. (Eds.) (2013). Tasks of the European Union Science Olympiads 2003–2007. In: *Challenging Interdisciplinary Science Experiments*, Vol. 1, Münster, Waxmann.

Dillon, J., Osborne, J. (2008). *Science Education in Europe: Critical Reflections*. London, The Nuffield Foundation.

EACEA/Eurydice, Eurostat (2009). *Key Data on Education in Europe 2009*. Brussels, Eurydice.

European Commission. (2007). *Science Education Now: A Renewed Pedagogy for the Future of Europe*. Brussels, European Commission.

Gray, D. (2006). *Doing Research in the Real World*. London, SAGE Group.

Hayes, D., Mills, M., Christie, P., Lingard, B. (2006). *Teachers and Schooling*. Sydney, Allen and Unwin.

Kemmis, S., McTaggart, R. (2003). Participatory action research. In: Denzin, N., Lincon, Y. (Eds.) *Strategies of Qualitative Inquiry*. London, Sage.

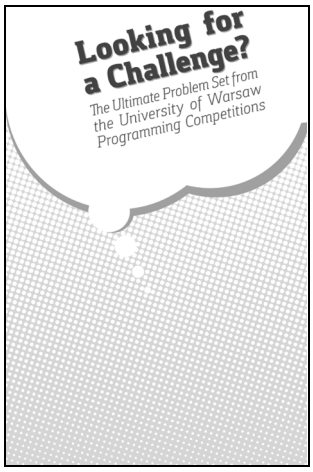
Lather, P. (1986). Research as Praxis. *Harvard Educational Review*, 56(3), 157–277.

OECD (2009). PISA 2009. *Assessment Framework – Key Competencies in Reading, Mathematics and Science*. Paris, OECD.

Stringer, E.T. (1999). *Action Research*, 2nd edn. Thousand Oaks, CA, Sage.

Dr. Carmel Mulcahy

Olympiads in Informatics, 2013, Vol. 7, 191–192



Looking for a Challenge

Editors: Krzysztof Diks, Tomasz Idziaszek, Jakub Łącki, Jakub Radoszewski
Publishing house: Faculty of Mathematics, Informatics and Mechanics, University of Warsaw
Country: Poland
Year of edition: 2012
Language: English
Number of pages: 428

One of the constants during my involvement with programming contests is the fact that there are always good problems from Polish contests. Problems from X and XI OI (2003 and 2004 Polish IOI team selections) were among the first and most memorable hard problems shown to me by my mentors. Most of my conversations at IOIs with Chinese team members inevitably drifted towards how we solved that year's POI problems. As a coach, a significant portion of the problems that I point contestants towards are from POI/PA/AMPPZ. The book 'Looking for a Challenge' is a compilation of the best among these problems, and will be a valuable resource for contestants at all levels for at least the next decade.

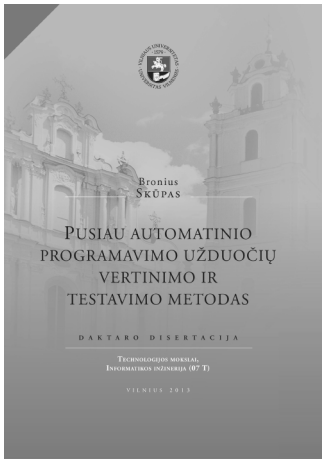
The book is organized into a series of disjoint chapters, each about a problem and its solution. Each set of chapters was written separately by a long time contributor to Polish programming competitions. These authors include the founders of the Polish contest programs as well as top contestants throughout the years. As the problems were chosen separately, the solution analyses often reflect the contributors' unique problem solving approaches. These insights are perhaps even more valuable than the problems themselves. They represent a large sample of top problem solvers' way of approaching, breaking down, and solving a problem. The problems selected cover a wide range of difficulty levels. Problems such as Canoes (p. 338), Triangles (p. 352), and Party (p. 395) are solved by elementary but tricky observations. On the other end of the spectrum, some problems' solutions naturally lead to key ideas from active areas of CS research. For example,

Ants (p. 18), Superknight (p. 240), and Guesswork (p. 416) are related to streaming algorithms, lattice reduction, and Bayesian statistics respectively. In between, many of the problems such as Permutations (p. 91) and Cave (p. 264) represent recurring themes in Polish contests with many interesting follow-ups and variations.

The organization of the chapters makes this book more suitable as supplemental material for developing problem solving skills. In this role, the choice of problems once again shines through. Many solutions have short implementations, making them ideal for students who are just getting familiar with programming. Each chapter includes a link to an online judge where readers can test if their solutions are correct. Most of the solutions are written in a self-contained manner, but as one gain more exposure to various techniques, the analysis can also be read once again from a more systematic perspective. Perhaps a good approach to using this book is to treat it as a series of disjoint articles, each to be read individually but repeatedly over time. When I heard that a friend from Warsaw University was visiting CMU, I suggested that he replace all empty space in his luggage with copies of 'Looking for a Challenge'. It is with the same enthusiasm that I recommend this book to you. It is a wonderful resource that provides a glimpse of the past, present, and future of Polish programming competitions.

Richard Peng

Olympiads in Informatics, 2013, Vol. 7, 193–197



Method for Semi-Automatic Evaluation and Testing of Programming Assignments

Abstract of dissertation in English, 33 pages

Author: Bronius SKŪPAS
Publishing house: Vilnius University
Country: Lithuania
Year of edition: 2013
Language: Lithuanian
Number of pages: 110

This is a PhD thesis (dissertation) defended at Vilnius University on February 11, 2013. The reason for this review in the journal is that it describes the developed system for evaluation of programming assignments, which is very important while teaching programming in secondary schools as well as training students for informatics olympiads.

Teaching programming is a difficult process because of the association with creative thinking, strictly formalized tasks, and practical programming assignments. The students have to be trained to create not only running, but also qualitatively designed reliable, properly functioning programs. Testing and evaluation of programs developed by the students requires a lot of the teacher's efforts and time.

In the thesis, automatic program testing is defined as dynamic testing, based on black box testing with tests prepared in advance. In order to use automatic program testing, programming tasks are often specified in detail, i.e., the required data input and output formats are introduced. This allows evaluating by applying the fact-verification method. Accurately identified verifiable fact is defined as an evaluation criterion.

Automatic programming assignments testing can be used in the teaching process, during programming exams, professional programming knowledge tests during recruitment and programming competitions.

Many researchers note that fully automated testing, based on static and dynamic analysis, cannot be completely fair. Therefore common practice is to use semi-automated testing of programming assignments, which is a mixture of automatic testing and manual evaluation and provides greater flexibility in the use of automated testing benefits.

The thesis focuses on the exploration of possibilities of using a semi-automatic testing system for programming assignments in schools, distance education, as well as competitions and exams. Common evaluation errors are analysed as well. The research is focused on creating a new method of evaluation for programming assignments in order to achieve high quality evaluation in acceptable time and the justification of the evaluation to the user.

The aim of the work was to develop a method for semi-automatic evaluation and testing of programming assignments that would improve system and evaluator interaction in order to increase the efficiency and the quality of the evaluation and to implement it into a prototype of semi-automatic evaluation system.

The constructive research method was chosen as the key method for the research. Kasanen *et al.* (1993) states six phases of the constructive approach: (1) finding a practically relevant problem with research potential; (2) obtaining general and comprehensive understanding of the topic; (3) innovating, i.e., constructing a solution idea; (4) demonstrating that the solution works; (5) showing the theoretical connections and the research contribution of the solution concept; (6) examining the scope of applicability of the solution.

When designing the Maturity exam of Information Technology/Informatics in Lithuania it was decided that the submitted programs designed by graduating high school students should be evaluated even if they do not compile. It was also agreed that the evaluation should be positive, i.e., the points should only be given for the skills demonstrated by the student. From the exam requirements it was clear that evaluation should be either manual or semi-automated. Manual evaluation was excluded due to limited number of evaluators and the limited time resources.

When considering the possibilities of applying semi-automated evaluation for evaluation of the Maturity exam submissions, a raised hypothesis was that deeper analysis of the process of manual evaluation of programs designed by the students would enable the creation of a new semi-automated testing and evaluation method which would have the same quality and precision as manual evaluation but higher efficiency (Fig. 1).

Long observation of informatics teachers evaluating programs manually as well as observation of the students searching for mistakes in their programs served as a premise for stating the hypothesis above. The process of designing a short program (i.e., no longer than 100 lines) is similar to that of designing complicated software systems.

The researcher attempted to benefit from the experience of evaluation in the Lithuanian informatics olympiads. However it turned out that the demands of the exam are different from that of the olympiads:

1. The evaluation system should operate in the environment as similar as possible to that used during the exams by the students (Windows OS, FreePascal compiler; after 2010 also Mingw g++).
2. The task cannot be considered as partially solved if the program passes just some of the grading tests. The tasks are rather simple and the reasons for each failure should be identified.
3. The program with minor mistakes (missing punctuation mark, undeclared variable, etc.) cannot be assigned few or no points because of those mistakes only.

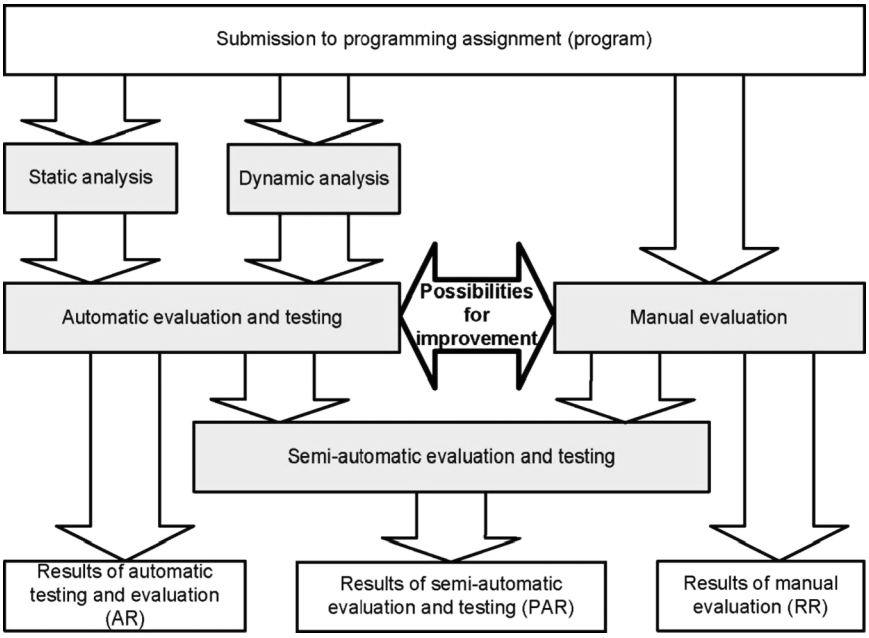


Fig. 1. Scheme illustrating suggestions how to improve the evaluation process.

4. There is no need to support the role of the student in the exam evaluation system, because the student cannot observe his/her score directly during the exam.

During the evaluation the semi-automated evaluation system should provide the administrator functionality to:

- load task specifications together with tests into the system;
- modify task evaluation criteria;
- to compose, assign and present to the evaluator an evaluation package containing a set of submitted programs to be evaluated;
- to assign an evaluation package to another evaluator;
- to observe and record the outcome of evaluation process.

A semi-automated evaluation system should provide the following functionalities for evaluating a submitted program:

- to be able to execute statistical analysis plug-ins and provide the results to the evaluator;
- to automatically test the submitted program using black box testing;
- to present the results of automated testing and the program being evaluated to the evaluator;
- allow modification of the program being evaluated and to retest the modified program without changing scores assigned by automated testing; to display the successfully passed tests, to visualize the changes made by the evaluator and to allow to restore the initial version of the submitted program;

- allowing to fill/modify the scores obtained from semi-automated testing if automated testing did not assign any points to the program;
- allowing to fill/modify the manual evaluation scores.

There were also added *non-functional requirements*. Those will not be emphasized in the formalized problem statement in this thesis, because their implementation is not so abstract. Here are the most important ones:

Semi-automated evaluation system should allow:

- to evaluate incomplete programs in a positive manner;
- to help finding the exact location of an error;
- to provide clear feedback messages to the user-evaluator;
- to operate fast;
- to allow the evaluator to comfortably experiment with the program being evaluated.

After formalizing the problem and emphasizing the role of the evaluator (i.e., dissociating from many roles common in various grading systems) the prototype of semi-automated evaluation system was created, which implemented the grading of one program. In the system there was implemented safe black box testing, reviewing the program under evaluation and collecting scores from manual grading.

After some experiments with the evaluation system being developed it was observed that for the evaluators it is not easy to find a mistake in the program. In such case the evaluators used to open the program in some widely accepted IDE and started experimenting there. However most IDEs do not provide functionalities for testing the program with a group of tests. This observation suggested the idea to transfer some IDE functionalities to the evaluation system. The first and the essential improvement was the following: the program text review component was replaced by program text editing component and there was added a button allowing batch retesting of the modified program. This improvement is illustrated in Fig. 2.

In conclusion, the researcher investigated and classified the functionalities of modern automatic and semi-automatic programming assignments testing systems and discovered that this class of software still encounters a lot of problems, the systems are being improved and new testing systems are being developed.

Improvements to the semi-automatic evaluation method presented in this work allows the evaluator to use automatic testing system interactively, to carry out the experiments with the program being evaluated, to analyse interactively the consequences of the errors discovered in the programs of the students, and to consider the weights of the mistakes made by the students.

The Lithuanian State Matura exam of Information Technology practical task assessment software presented in this work implements the proposed semi-automatic evaluation method. The system has been used successfully by the National Examination Centre; it is constantly being improved, adapted to the changing needs.

The effectiveness of manual evaluation method was compared to the proposed semi-automated evaluation method: the efficiency increased from 1.4 to 2.2 times.

Qualitative analysis of evaluation results shows that results obtained by using the proposed semi-automatic evaluation method and the automatic testing differ significantly.

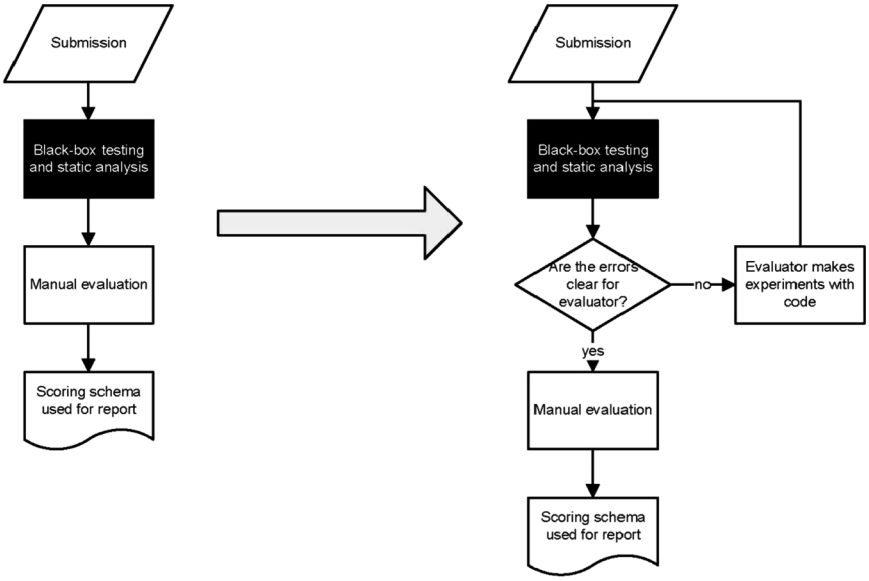


Fig. 2. Semi-automated program testing process improvement.

It was confirmed that the application of the proposed method resulted in higher quality of evaluation.

The method has also been also been implemented and tested in the EduJudge plug-in (`http://eduvalab.uva.es/en/projects/edujudge-project`) to the learning management system *Moodle*. (It is pity that this plugin is still not available for general public). The experiment confirmed that the improved semi-automatic evaluation method can be easily implemented in other systems.

References

Kasanen, E., Lukka, K., Siitonen, A. (1993). The constructive approach in management accounting research. *Journal of Management Accounting Research*, 5, 243–264.

Valentina Dagienė and Bronius Skūpas
(prepared by Abstract of dissertation)