2018

CONTENT

ERIES JOURNAL

JOURNAL ON EFFICIENCY AND RESPONSIBILITY IN EDUCATION AND SCIENCE

VOLUME 11

ISSUE 4

Web pages: http://www.eriesjournal.com

Scientific journal of the Czech University of Life Sciences Prague JOURNAL ON EFFICIENCY AND RESPONSIBILITY IN EDUCATION AND SCIENCE, distributed by the Faculty of Economics and Management. Published quarterly. Executive editors: Ing. Martin Flégl, Ph.D. and Ing. Igor Krejčí, Ph.D., Editorial Office**:** ERIES Journal, Czech University of Life Sciences Prague, CZ 165 21 Prague 6 - Suchdol, Czech Republic, email: editor@eriesjournal.com, tel: +420 224 382 355.

CZECH UNIVERSITY OF LIFE SCIENCES PRAGUE

An international peer-reviewed journal published by
Faculty of Economics and Management
Czech University of Life Sciences Prague

contact: editor@eriesjournal.com
www.eriesjournal.com
Online ISSN: 1803-1617
Printed ISSN: 2336-2375

EDITORIAL BOARD

Editorial Office

ERIES Journal, Czech University of Life Sciences Prague, CZ 165 21 Prague 6 - Suchdol, Czech Republic,
email: editor@eriesjournal.com,
tel: +420 22438 2355

Registration number: MK ČR E 21414

INSTRUCTIONS FOR AUTHORS

The Journal on Efficiency and Responsibility in Education and Science publishes papers of the following categories: full research papers, short communications, review studies and book reviews (on invitation only).

- FULL RESEARCH PAPERS
- SHORT COMMUNICATION
- REVIEW STUDY

Papers are published in English. A paper may comprise an empirical study using an acceptable research strategy, such as survey, case study, experiment, archival analysis, etc. It may contain a theoretical study aimed at advancing current theory or adapting theory to local conditions or it may arise from theoretical studies aimed at reviewing and/or synthesizing existing theory. Concepts and underlying principles should be emphasized, with enough background information to orient any reader who is not a specialist in the particular subject area.

Submission checklist

The paper. The paper is carefully formatted according to the template of the journal (see bellow). Special attention is paid to the exact application of the Harvard referencing convention to both continuous citations and list of references. If an electronic source has the DOI number assigned, also it will be provided in the list of references. Manuscripts are submitted via the editorial system in the DOC.

Research highlights. The core results, findings or conclusions of the paper are emphasized in 1-3 bullet points (max. 100 characters per bullet point including spaces). The highlights are submitted as a text into the submission form in the editorial system.

Copyright form. The submission of a paper will imply that, if accepted for publication, it will not be published elsewhere in the same form, in any language, without the consent of the Publisher. The manuscript submitted is accompanied by the copyright form signed by the corresponding author who declares the agreement of all authors with the conditions in the Form. The Form is submitted into the editorial system in the PDF format.

Suggested reviewers. It is required to suggest two experts appropriate to evaluation of the paper. The experts should be out of the affiliation of the author(s), Czech University of Life Sciences Prague, and also both experts should be from different affiliations. The reviewers are submitted into the text fields in the submission form of the editorial system.

Preparation of the manuscript (technical notes)

Authors are responsible for applying all requirements that are specified in the journal's paper template in individual sections. Especially, the paper must provide a short review of current state in the area of the paper's aim in Introduction. The paper should refer significant sources, particularly scientific journals or monographs.
Papers must be closely scrutinized for typographical and grammatical errors. If English is not author's first language then the paper should be proof-read by a native English-speaking person, preferably one with experience of writing for academic use. Spelling should follow the Oxford English Dictionary.
Tables, graphs and illustrations should be drawn using a suitable drawing package. Colour may be used. Place all diagrams and tables where you wish them to appear in the paper. Ensure your diagrams fit within the margins and are resizable without distortion.

Review procedure

Following Editorial recommendation, papers are submitted to a double-blind peer review process before publication. Commentary by reviewers will be summarized and sent by email to authors, who can choose to revise their papers in line with these remarks. Re-submitted papers should be accompanied by the description of the changes and other responses to reviewers' comments (see above), so that the desk-editor can easily see where changes have been made.

Copyright

Authors are fully responsible for the paper's originality and for correctness of its subject-matter, language and formal attributes. Author's statement should be enclosed declaring that the paper has not been published anywhere else.

The submission of a paper will imply that, if accepted for publication, it will not be published elsewhere in the same form, in any language, without the consent of the Publisher. Before publication, authors will be asked to complete a copyright release, giving the publisher permission to publish the paper in a specific issue of this Journal. Overall copyright ownership of the paper, however, remains with the author/s. It is the authors' responsibility to obtain written permission to quote material that has appeared in another publication.

EDITORIAL

With this issue you are holding in your hands, Journal on Efficiency and Responsibility in Education and Science (ERIES Journal) has finished its first year from the second decade of its existence. In this last issue of the year 2018 (vol. 11, no. 4) we have a variety of articles with a centre topic of students' motivation and reasoning. It is of a high importance of universities to offer study programs that attract students' interests to learn. On the other hand, students' motivation is one of the essential factors that propels learning. In most of the cases, students prefer learning directly related to real cases presented by experienced professors. Students must feel that the education they receive has practical applications. As students may say, they want an education that serves to their needs. To propel students' learning, there should be a synergy between the quality of teaching staff and motivated students. This synergy can lead to higher quality in education, as well as in research, which can then generate new ideas and findings. To do so, students, as well as professors, must be willing to step out from their comfort zones and push themselves towards new unrecognized limits. Students should be taught to think in a different, uncommon way, which stimulates their creativity. Students must find their way to discover solutions in their problem-solving.

The first article from authors Kateřina Berková, Jana Borůvková and Lenka Lízalová present an analysis of factors that influence students' motivation to study economic subjects in economic and technical study programmes. For this purpose, the authors use a sample containing 82 respondents from Finance and Management and Applied Computer Science study programs. Collection of data was performed using an online questionnaire through an e-learning platform Moodle. Results reveal that it is generally simpler to motivate students of technical study programmes since these students are more interested in their study programmes. What is more, students of technical study programmes are more demanding and have higher expectations related to their study program compare to students of economic study programs. The authors conclude that to support students' motivation, it is necessary to provide experience-focused learning based on an active problem-solving process in the form of live stories comprising a subsequent analysis of mistakes and guidelines for reviewing the respective problem.

The second article from Mikuláš Gangur and Miroslav Plevný presents a possible way of solving the problem of creating more test variants for a large number of students divided into groups. The similar difficulty of each variant is considered as a key issue of test variants design. The authors introduce an automatic parametrized test generator. The article describes the functionality and implementation process of the generator stemmed from the set fields by means of a combination of questions according to the set criteria. The proposed generator decreases the complexity computing with respect to increasing number of demanded questions. The results show that the time of generation of the test was shortened from hours to seconds in complex cases. This allows teachers to use the generator in real time.

The last article from Jarmila Novotná and Martin Chvál investigates how the order of numerical data in word problems on division of a whole into unequal parts affects achievement and reasoning of 14-16-year old students. The analysed sample consists of 182 grade 9 students from four Czech primary schools, where all schools were with no specialisation, of medium size and attended by children from their immediate surroundings with a varied socio-economic background. Students from the sample were divided into equally skilled groups. These groups were later given distinct versions of the same word problem that differ in the order of numerical data in the problem statement or in the context or the presence of "if-clause". The results indicate that students' success is affected by the order of numerical data in the statement in an unfamiliar context. Understanding the impact of the order of data on the difficulty of problems is important not only for teachers but also for authors of tests, curricular documents and textbooks.

We hope that all our readers will find this last issue of the year 2018 interesting. We also hope that ERIES Journal will contribute to the field of efficiency and responsibility in education as it has contributed so far. With the end of the year 2018, we would like to thank all the authors who have submitted their manuscripts to ERIES Journal, to all reviewers who carefully reviewed all these manuscripts, and to all members of the Editorial board who contributed in increasing the ERIES Journal quality.

We wish you Merry Christmas and all the best in 2019.

Sincerely

Martin Flégl
Executive Editor
ERIES Journal

https://www.eriesjournal.com
https://www.linkedin.com/company/eriesjournal/
https://erie.pef.czu.cz

MOTIVATION OF STUDENTS OF ECONOMIC AND TECHNICAL STUDY PROGRAMMES AS A TOOL OF COMPETITIVENESS OF UNIVERSITIES AND COLLEGES: EMPIRICAL STUDY

Kateřina Berková[1✉], Jana Borůvková[2], Lenka Lízalová[2]

[✉1]College of Polytechnics Jihlava, Jihlava, Czech Republic, katerina.berkova@vspj.cz

[2]College of Polytechnics Jihlava, Jihlava, Czech Republic

Highlights

- *The increase of students' motivation proves to be essential for the competitiveness of universities*
- *Students are the most significantly motivated by experience-focused learning*
- *Students of the technical study programmes assess the motivation factors more strictly*

Abstract

The article presents an analysis of factors influencing the students´ motivation to study economic subjects in economic and technical study programmes in the conditions of a practically oriented college. The research is focused on quality management systems of universities and the competitiveness in the tertiary education market. Specifically, the area of quality of higher education is addressed, through students' assessment of the subject of Business Administration. Motivation factors related to teaching techniques were classified and the intensity of motivation of the respondents was investigated in the questionnaire survey. The statistical analysis of data suggests that there are differences in the motivation of students of economic and technical study programmes. Students of technical study programmes Applied Computer Science are more demanding and have higher expectations from study programmes compared to students in the study programmes of Finance and Management, whose expectations are lower. The results are discussed and compared with a similar foreign research. A pedagogical constructivism should be introduced into the Czech education, resulting in enhancement of attractiveness of the study programmes, thus increasing the level of competitiveness of the college.

Article type

Full research paper

Article history

Received: September 21, 2018

Received in revised form: November 17, 2018

Accepted: December 6, 2018

Available on-line: January 1, 2019

Keywords

Motivation of students, economics, technics, universities' competitiveness, colleges' competitiveness

Berková K., Borůvková J., Lízalová L. (2018) "Motivation of Students of Economic and Technical Study Programmes as a Tool of Competitiveness of Universities and Colleges: Empirical Study", *Journal on Efficiency and Responsibility in Education and Science, Vol. 11, No. 4*, pp. 72-77, online ISSN 1803-1617, printed ISSN 2336-2375, doi: 10.7160/eriesj.2018.110401.

Introduction

Competitiveness of universities represents a reflection of the quality of each individual institution, which is based on its internal evaluation. Nowadays, the quality has become one of the competitiveness tools on the market of tertiary education. Different approaches are used in the Czech Republic - universities primarily utilise models that have become well-established commercially or on the European level such as Common assessment framework (CAF) and the European foundation for quality management (EFQM). Furthermore, they use the Total quality management (TQM) model, which is based on the principle of positive motivation and focuses on the quality of stakeholders and all resources; in addition, ISO certification of 9000 series is used (Krpálková Krelová, 2014; Mulač and Mulačová, 2014). Measurements of quality and evaluation are also possible through assessment of teaching programmes, i.e. through students who co-participate in the evaluation process (Mulač and Mulačová, 2014).

Currently, the universities have been experiencing a decrease in the number of students, to which they attempt to respond using various tools, in accordance with the strategic objectives and quality assurance systems implemented by universities. One of the target tools is monitoring and reinforcement of the motivational potential, i.e. increasing of the attractiveness of the study programme, which increases the potential of obtaining students. This increases competitiveness in the market of tertiary education. As mentioned above, the problem is addressed by, for example, Savenkova and Olesina (2018) or Kim and Regh (2018). The study-oriented motivation of students has become increasingly significant and, considering the conditions of decreasing numbers of students in the Czech Republic, it becomes one of the predominant tools of competitiveness at universities. The above-mentioned facts prove that the matter of study-oriented motivation of students is a rather essential variable in the quality assurance system at universities.

Motivation plays a significant role in the branch of education. Fontana (2014: 153) claims that motivation is one of the fundamental effects which determines the efficiency of the educational process. There is, however, a high number of factors that affects the intensity of motivation. In general, they comprise mainly the teacher's effect on students, i.e. his/her professional as well as overall competence profile (Pasiar et al, 2015). Nevertheless, it is not possible to disregard the content of the study programme as such (the subject or the whole branch), self-strictness of students and their expectations regarded to the study programme (Berková, Novák and Pasiar, 2018). Expectations of individual persons and their fulfilment may be regarded as an important factor shaping the level of intensity of motivation. This fact results from the classification of motivation, which is defined as a summary of strong suites that organize the behaviour and experience of an individual with the objective of changing the existing unsatisfactory situation or achieving something positive (Plháková, 2003: 319).

In the present world of technological progress, which allows for a different manner of education – in a more interactive format – it is increasingly difficult to attract students and motivate them to learn. Students may find information themselves and quite easily as it is accessible through modern resources. It is essential to ensure that the teacher is motivated to maintain the development of his/her professional and pedagogical competences. The aforementioned motivation is reflected in an increased level of students' motivation (Stranovská, Lalinská and Boboňová, 2018). Findik-Coskuncay, Alkis and Ozkan-Yildirim (2018), publish a relatively new research focused on management systems of the quality of learning of 470 university students through improvements in e-learning services.

The objective of this study is to determine factors that affect behaviour in the area of tertiary education and learning management systems. The proposed constructional model was verified on the basis of modelling of structural equations. Using these methods, it was determined that students perceive e-learning positively as regards its efficiency, ease of use, pleasure and satisfaction with the interactive environment. Furthermore, they made a positive evaluation of the feedback which e-learning provides throughout the study programme. Interactive education has become a widespread format of the process of teaching and learning. This element implemented through modern technologies, communication technologies, allows for real-time education based on interaction. It further increases the learning-oriented motivation of students, may increase the active participation of students and positively influence the process of learning in the branch of tertiary education (Tan et al, 2018). E-learning, considering its sophistication and complexity, represents an educational platform suitable for the university environment (Ogrodzka-Mazur et al, 2017). It allows for the development of higher cognitive processes such as analysis, evaluation and creativity, which require an extensive interaction and collaboration between students and facilitators (Lau et al, 2018).

Currently, the role of students is gradually transforming into the role of clients of universities in the Czech Republic. Therefore, students no longer value their education and services provided to them by universities. This fact results, inter alia, from a demographical decrease which is currently reflected in tertiary education. The breakpoint occurred in 2012-2014, while around 2023-2025 the demographic curve is expected to be increasing in tertiary education once again (Koucký and Bartušek, 2011). The decrease of the number of students might (even unintentionally) cause decreases in the quality of the educational process, which represents one of the variables of quality monitoring at universities. Mulač and Mulačová (2014) stipulate that financial allocations to individual universities are based on their quality. The aforementioned fact is also confirmed in the study drafted by Šipikal and Némethová (2017: 170), who address the matter of financing of universities on the basis of their performance. They emphasise the importance of objectives of public policies – increasing the participation of universities in education, international interaction and cooperation with the practice, etc., which should be reflected in the calculation of the performance of each respective university.

The content of education reflecting trends from the practical application of economics as well as modern means and teaching methods are definitely the factors which increase attractiveness of a study branch, respectively individual subjects. This also leads to an increase in the motivational potential of the study programme. Nevertheless, it is to be mentioned that students' motivation is significantly determined by their teacher's personality. The level of the teacher's pedagogical competences is based on the qualifications, scope, form and quality of practical skills, specialisations, age, professional experience, verbal competences and attitudes (Windham, 1988). An ideal teacher features resistance, adaptability, ability to absorb new knowledge, social empathy and communication skills (Dytrtová and Krhutová, 2009: 15). A suitable form of teaching has positive effects on education. Students featuring insufficient levels of motivation are not interested in the course of solving of the respective subject, they are not concerned about mistakes which they make or correct procedures; they are only interested in the final result of the task (Boekaerts, 2004). The research of Berková and Krejčová (2016), which inspired the above-mentioned new research, shows that students perceive expert knowledge of the teacher as the most significant motivational factor.

Considering the fact that the applied teaching methods and style of pedagogical work of the teacher as well as the teacher's motivation and professional competences affect students' motivation (Brožová, Horáková and Fiedler, 2018), the respective matter was continuously solved within the scope of a systematic research lasting several years. In their research, the authors (Berková, Králová and Krejčová, 2017; Berková and Krejčová, 2016; Krejčová and Berková, 2016) dealt with the matter of utilisation of modern teaching methods, i.e. experience-focused learning and teaching on the basis of pedagogical constructivism in the study programmes of economics. The most commonly used methods are the discussion methods followed by methods of problem-focused learning such as problem-oriented questions; case study methods prove to be on a significant level as well. In conclusion, the authors state that in the Czech environment modern methods are still not as commonly used as abroad. Nowadays, teaching on the basis of interactive environments utilise information technologies prove to be rather neglected. In addition, one of the topics which are currently discussed is the type of teachers which should be trained for the new generations of students, and which subjects should be implemented with the objective of reinforcement of students' motivation as regards learning and obtaining general competences (Zygaitiene and Buivydaite, 2017, compared to Au, 2011).

Research Problem and Objective

The research is focused on reviewing the factors which affect students' motivation to study the subject Economics. Specifically in case of students studying economic and technical study programmes in conditions of a non-university type of the College of Polytechnics Jihlava, Czech Republic. This is a primary research, as the research sample will be extended to other relevant college students prospectively. The results will be further used to experimentally measure the effects of experiential methods on professional competence. Considering the fact that students of the study programmes Finance and Management and Applied Computer Science study the subject Economics together, the present study is focused on the motivation of students in relation to studying of this subject. The motivation factors are related solely to the teaching level and a comparison of the intensity of motivation between the branches is presented. Another objective of the study pertains to the determination of factors which students assess similarly and those which they assess completely differently, separately for students of Finance and Management and also separately for students of Applied Computer Science. The subject of the research is the research hypothesis that the perception of motivation factors by students of economic and technical study programmes in connection with studying of the subject Economics is different between the study programmes as well as inside each of the study programmes. This research was prompted by the current decrease in the number of students in tertiary education; its contribution is in the final recommendation drafted on the basis of an empirical analysis. The article presents inspiring tools for the increase of students' interest in studying economics.

Materials and Methods

The research was conducted concurrently for both study programmes, within the scope of teaching in the term of February-March 2018. Students of the first year of Finance and Management (n = 52) and students for the first and second year of Applied Computer Science (n = 30) at the College of Polytechnics Jihlava in the Czech Republic, a non-university type of school, participated in the research. In total, the research sample comprised 82 respondents, the selection of whom was intentional. The principal criterion of selection was a comparison of the intensity of motivation of students studying economic and technical study programmes as regards studying of the subject of Corporate Economy immediately after its completion, in conditions of a non-university type of college. Considering the fact that the students had completed the aforementioned subject in the preceding winter semester of 2017/2018, it was desirable to include them in the selection.

The factors in case of which the students assessed motivational effects were divided into three sets; each category comprised three selections:

1. **Form of elaboration of the educational content** (F1 Implications of the teaching materials; F2 Reference examples for preparation for credit test/examinations; F3 PowerPoint presentations, e-learning support, professional literature without extensive relevance);
2. **Maintaining of students' attention in class** (F4 Discussion, questions during the lecture and time for enquiries; F5 Stories from the practice, case studies and solving of problems from the practice; F6 Independent work of the students related to a task and the teacher's corrections);
3. **Style of explanation of lessons** (F7 Teacher's guidance how to think – analogies, critical analysis and evaluation of information; F8 Acceptance of less frequent ideas presented by students and analysis of mistakes; F9 Teacher presents more than just facts, the student is able to coordinate their cognitive resources, emotions and actions in relation to learning goals, the student acquires metacognitive experience, realizes his own cognitive ability).

Selection of the factors was inspired by a long-term research conducted in 2015-2017 which was focused on the motivation of students at secondary schools of economics, in relation to studying of economic subjects (Pasiar et al, 2015; Berková and Krejčová, 2016; Berková et al, 2018). These factors were also chosen for the purpose of comparing the results of motivational perception of the factors by students of secondary and tertiary education as regards future development of solving of the research task. Furthermore, they were selected considering the primary objective of our research, i.e. ensuring motivation of students of economic and technical study programmes in conditions of a non-university type of college, with the principal focus being placed on the level of teaching (teacher's lectures) and students' learning.

Collection of data was performed using a questionnaire which was drafted in the online e-learning environment of Moodle. The structure of the questionnaire was identical for both reviewed study programmes. The questionnaire contained 2 questions determining factual data (gender and study programmes) as well as 9 factors in case of which the students were to evaluate the intensity of motivational potential on the scale between – to 2. (The lowest levels of intensity of motivational potential, i.e. –2 and –1, represent a demotivating effect of the factors while the highest levels of intensity of motivational potential, i.e. the values of 1 and 2, represent a motivating effect of the factors; the value of 0 represents the respective student's indifferent relation to the factor.)

The non-parametric Mann-Whitney U test was used. Selection of non-parametric statistics proves to be justifiable considering the values of compared variables. They represent numeric, ordinal variables which comprise max. 5 categories on the scale between –2 and 2. Furthermore, correlation coefficients were calculated for all pairs of motivation factors, with the objective of determining which factors students evaluate similarly and which factors students evaluate completely differently – once again separately for students of Finance and Management and students of Applied Computer Science. Kendall's tau-b was selected from the high number of existing correlation coefficients as the analysed variables were ordinal and featured the same scale. The respective coefficient ranges between –1 and +1. If it approximates zero (it is not statistically significant), the factors are independent – students evaluate the factors independently (differently).

Null (statistical) hypotheses, which form the subject matter of verification on the 5% level of significance, are following:

H_{0-1}: *Mean values of the motivation factor are identical in both study programmes.*

H_{0-2}: *Students of Finance and Management evaluate all the motivation factors identically.*

H_{0-3}: *Students of Applied Computer Science evaluate all the motivation factors identically.*

Results

Comparison of Motivation Factors

The resulting order of factors for Finance and Management (FM) n = 52 and Applied Computer Science (ACS) n = 30 is presented in Table 1.

Motivation Factors	Mean		Median	
	FM	ACS	FM	ACS
Examples and stories from the practice	1.31	1.33	1.50	1.00
Teacher presents more than just facts	1.19	0.97	1.00	1.00
Implications of teaching materials	1.15	1.13	1.00	1.00
Analysis of mistakes, acceptance of less frequent ideas	1.12	1.37	1.00	1.50
Teacher's guidance on how to think	1.13	0.90	1.00	1.00
Questions, discussion, enquiries during lecture	0.54	0.73	1.00	1.00
Reference examples for exams, without practice	0.37	–0.13	1.00	1.00
PPT presentations, e-learning, books	0.02	–0.47	0.00	0.00
Independent work in lessons	–0.12	–0.97	0.00	0.00

Table 1: Descriptive Statistics: Motivation Factors in the Study Programmes of Finance and Management, and Applied Computer Science, 2018 (source: own calculation)

The scope of significant factors for students of Finance and Management comprises – considering the determined mean values – any and all factors from the category of *Form of elaboration of the educational content*. This category consists of following factors from table 1: Teacher's guidance on how to think; Analysis of mistakes, acceptance of less frequent ideas; Teacher presents more than just facts. This last factor means that the student is able to coordinate their cognitive resources, emotions and actions in relation to learning goals, the student acquires metacognitive experience, realizes his own cognitive ability. The examples and stories from the practice and solving of problems are also included in the category of *Form of elaboration of the educational content*. This factor corresponds to the factor of Implications of teaching materials. The intensity of the aforementioned factors exceeds the value of 1.

In addition, the results clearly show that students feel demotivating effects of independent work in lessons in case of which the teacher solely reviews the correctness of their solutions without any in-depth connotations or discussions. This is also related to the fact that e-learning support, PowerPoint presentations and professional literature are not motivating for students either, respectively they are indifferent to this factor. If such materials were not available, their motivation related to studying economic subjects would not be affected. The motivation factor of Questions and enquiries during lecture, in case of which the value is higher (0.54), may be interpreted similarly. Current generation of students in the first year of Finance and Management is most motivated by examples and stories from the practice, on which it is essential to focus more closely, thus focusing on problem-solving tasks featuring attractive elements.

Students of Applied Computer Science assess the reviewed factors more critically and strictly than students of Finance and Management. This fact is the most obvious in case of the factors of Reference examples for exams, without practice; PPT presentations, e-learning, books; and Independent work in lessons. The mean values representing the intensity of the motivational effects are negative, i.e. these factors are demotivating for the students – unlike the indifferent relation to these factors determined in case of students or Finance and Management. Students of Applied Computer Science would be most motivated by Stories and examples from the practice; Analysis of mistakes, acceptance of less frequent ideas; and Implications of teaching materials. In case of students of Finance and Management, the value of 1 was exceeded for more factors than in case of students of Applied Computer Science. In addition, students of the economic study programmes used negative values for fewer factors. This fact suggests that technical students of a practically oriented college have higher expectations from their college study programme as well as higher requirements related to the study programme. They are not satisfied with superficial knowledge.

Interdisciplinary Differences Related to Perception of the Motivation Factors

Table 2 presents *p*-values for all the motivation factors.

Motivation Factors	*p*-value
Implications of teaching materials	0.648
Reference examples for exams, without practice	0.059
PPT presentations, e-learning, books	0.080
Questions, discussion, enquiries during lecture	0.425
Examples and stories from the practice	0.662
Independent work in lessons	0.002
Teacher's guidance on how to think	0.308
Analysis of mistakes, acceptance of less frequent ideas	0.107
Teacher presents more than just facts	0.356

Table 2: *P*-values of Mann-Whitney Test – Comparison of Interdisciplinary Differences, 2018 (source: own calculation)

At the 5% level of significance, we reject the null hypothesis H_{0-1} solely in case of the factors of Independent work in lessons and Teacher's corrections of solutions.

The reviewed study programmes prove to differ only in case of the respective factor, whereas students of Applied Computer Science assess this factor more strictly and it is more demotivating for them than for students of Finance and Management. The other motivation factors do not feature statistically significantly different medians for the study programmes. It is to be emphasized that even though a statistically significant difference had not been proven, the study programmes show only minor differences in case of the following factors: Reference examples for exams/credit test, without practice; and Studying using e-learning support, books and PowerPoint presentations. The values of the above-mentioned factors are negative in case of students of Applied Computer Science while being positive in case of students of the economic study programmes (Table 1). On the basis of this fact it may be deduced that students of the technical study programmes are demotivated by these means while students of the economic study programmes prove to be indifferent to these factors, i.e. students of technical subjects appreciate analyses of procedures leading to solutions of each respective problem more than obtaining solely the final solution to the task (compare to Boekaerts, 2004).

Kendall's tau-b Correlation Coefficients – Pairs of Motivation Factors

The following pairs of motivation factors with a positive dependence have been determined on the 5% level of significance (students of Finance and Management perceive these pairs of factors similarly):

- **F1 Implications of the teaching materials** in combination with Discussion and questions during lectures (0.321); Teacher's guidance on how to think on the basis of analogies and critical analyses (0.258); Acceptance of less frequent ideas and analysis of mistakes (0.362); and Teacher's ability to present more than just facts (0.236);
- **F2 Reference examples for credit test/exam** in combination with PowerPoint presentations, e-learning support and professional books (0.404); and Teacher's ability to present more than just facts (0.226);
- **F3 PowerPoint presentations, e-learning support and professional books** in combination with Teacher's ability to present more than just facts (0.273);
- **F4 Discussion and questions during lessons** in combination with Stories and case studies from practice (0.252);
- **F5 Stories and examples from practice** in combination with Teacher's guidance how to think – analogies, critical analyses and evaluation of information (0.333); and Acceptance of less frequent ideas and analysis of mistakes (0.400);
- **F6 Students' independent work on a given task and the teacher's corrections** in combination with Teacher's guidance on how to think – analogies, critical analyses and evaluation of information (0.236).

At the 5% level of significance, we do not reject the null hypothesis H_{0-2} for the motivation factors.

The results document the fact that students of Finance and Management prove to evaluate similarly primarily factors within the same category, which might result from interconnections between them. They also perform similar evaluations of factors in case of which they might perceive similar purposes for use in the teaching process, even though they are from different categories. For example, Stories from practice; and Teacher's guidance how to think; respectively Discussion during lessons, in case of which it is obvious that real stories from practice invoke thinking and the teacher has to lead students to the manner in which they are supposed to solve each respective problem, which also requires discussion during the teaching process. In addition, it is possible to mention, for example, Stories from practice; and Acceptance of less frequent ideas and analysis of mistakes. Once again, in case of Stories and examples from practice, which prove to be focused on problems, the aforementioned manner of analysis of mistakes and acceptance of less frequent ideas of students is highly adequate. Therefore, students might perceive their motivation potential. However, the similar evaluation of the factors of Students' independent work on a given task; and Teacher's guidance how to think proves to be rather inconsistent with the above-mentioned results.

The evaluations presented by students of Applied Computer Science are stricter. It may even be claimed that the factors of Reference examples for preparation for credit test/exams, without practice; and Independent work in lessons have negative effects and teaching performed using these manners would not meet their expectations – in comparison with the students of Finance and Management. The latter students, on the contrary, evaluate these factors positively, even though the intensity of motivation proves to be lower. However, students of Applied Computer Science feature a similar approach to the importance of teaching materials comprising acceptance of less frequent ideas and analyses of mistakes; stories from the practice comprising acceptance of less frequent ideas as well as their teacher's guidance on how to think, with the teacher's property of presenting more than just facts.

Students evaluate differently the following two factors:

- F2 Reference examples for credit test/exam and F4 Discussion and questions during lessons (–0.217);
- F2 Reference examples for credit test/exam and F7 Teacher's guidance how to think
- (–0.260);
- F4 Discussion and questions during lessons and F6 Students' independent work on a given task and the teacher's corrections (–0.271);
- F5 Stories and examples from practice and F9 Teacher presents more than just facts (–0.345).

At the 5% level of significance, we reject the null hypothesis H_{0-3} for the factors F2xF4, F2xF7, F4xF6 and F5xF9.

Discussion

The research focused on the motivation of students in conditions of a practically oriented college confirmed the hypothesis that perception of motivation factors by students of economic and technical study programmes in relation to studying of economic subjects proves to be different on both the interdisciplinary level as well as within the scope of each study programmes as such. It resulted in determining of several fundamental facts which are to be approached materially while interpreting the respective relations with a certain amount of reservations.

As regards students of Finance and Management as well as students of Applied Computer Science, the significant factors which motivate them to study economic subjects include (a) Teacher's guidance on how to think; (b) Acceptance of less frequent ideas; and (c) Analysis of mistakes and Teacher presents more than just facts from the category of *Manner of explanation of lessons*. In addition, the scope comprised Examples and stories from the practice; and Looking for solutions to problems from the category of *Maintaining students' attention in class*. In order to prepare such stories, it is essential to utilise interdisciplinarity of individual subjects, interconnect individual findings and respect interdisciplinary relations between related disciplines on the level of subjects or topics (Kim and Regh, 2018; Savenkova and Olesina, 2018). Such principles, on the basis of which the teaching should be based (according to students), correspond to the nature and philosophy of experience-focused learning; however, this trend has not become so widespread in tertiary education yet (Berková, Novák and Pasiar, 2018). Nevertheless, it proves to be desirable – based on the students' requirements and expectations.

Students of the economic study programmes regard independent work during the teaching process, in case of which the teacher solely checks and corrects solutions without any in-depth connotations, as a demotivating factor. An indifferent relation was determined in case of the variables of discussion and studying based on e-learning support, PowerPoint presentations and professional books. The determined findings are, however, inconsistent with several recent foreign research papers which, on the contrary, prove attractiveness of e-learning being the interactive form of education (Findik-Coskuncay, Alkis and Ozkan-Yildirim, 2018; Lau et al, 2018; Tan et al, 2018). Students might regard these factors as demotivating because they require something else than their common experience at school. They require that the education should be more based on dialogue and, possibly, even more interactive. This means that e-learning support might not utilise its full potential while duplicating the standard, traditional text-based materials without any interactive elements or feedback.

Students of the technical study programmes (Applied Computer Science) assess the reviewed factors more strictly (critically), as determined in case of Reference examples for exams, without practice; Studying based on e-learning support, books and PowerPoint presentations; and Independent work in lessons (comp. Findik-Coskuncay, Alkis and Ozkan-Yildirim, 2018; Lau et al 2018; Tan et al, 2018). These factors prove to feature highly demotivating effects on this group of students.

It may be stated that significance of the factors proves to be assessed by students similarly within the scope of each reviewed study programmes of study, even though certain inconsistencies were determined as regards evaluation of the factors of Independent work in lessons and Teacher's guidance on how to think. It is rather interesting that students jointly compared teacher's guidance on how to think with the principles of experience-focused learning, i.e. on the basis of real stories and active learning. However, they also evaluated it similarly to the principles of traditional teaching without any attractive elements. The above-mentioned discrepancy might be explained by, for example, the fact that students perceive and interpret teacher's guidance on how to think differently. They might regard this activity as sole explanations of the study materials, respectively they are not led to becoming familiar with the significance of information during the thinking process (comp. Berková et al, 2018; Berková and Krejčová, 2016).

As mentioned above, the findings suggest possible manners of increasing the quality of education in accordance with the strategic objectives and quality assurance systems applicable to university education as well as support of acquiring students, as documented by certain national and international studies (Kim and Regh, 2018; Savenkova and Olesina, 2018; Mulač and Mulačová, 2014).

Conclusion

The objective of this study was to review the motivation factors applicable to students of economic and technical study programmes in relation to studying of economic subjects in conditions of a practically oriented college. Therefore, we emphasized the directions allowing for improvements in competitiveness on the market of tertiary education, in the area of improvement of quality of the educational process.

The present research determined that it is generally simpler to motivate students of technical study programmes in conditions of a practically oriented college since these students are really interested in their study programmes. Students of technical study programmes are more demanding and have higher expectations related to their study programme – in comparison with students of an economic study programmes (Finance and Management) whose expectations are lower. Therefore, it may be assumed that a teacher's motivational influence on a group of students of economic study programmes is a more complicated issue.

In spite of the above-mentioned principal differences in motivation of students of economic and technical study programmes at a practically oriented college, it may be stated that both groups would be most motivated by the so-called experience-focused learning based on an active problem-solving process in the form of live stories comprising a subsequent analysis of mistakes and guidelines for reviewing the respective problem. In order to maintain quality, it is desirable to include checks on the implementation of learning strategies in the form of evaluation tools. Although it is a primary research that has been placed in the conditions of a practically oriented college, in the area of increasing the motivation potential of economic subjects, the authors consider these recommendations in the university environment to be up to date. Teaching should be based more on experience. The authors will prospectively extend the research sample to other relevant college students. Further, authors plan to experimentally measure the effects of experiential methods on the development of professional competencies desirable in the labour market.

Acknowledgements

This research was supported by the College of Polytechnics, Jihlava, Czech under Grant No. 1170/4/181 "Model of Permeability of Corporate Economy and Financial Accounting in the Quality Improving System of the College of Polytechnics Jihlava".

References

Berková, K., Novák, J. and Pasiar, L. (2018) Modernizace ekonomického vzdělávání v kontextu taxonomií výukových cílů, Prostějov: Computer Media.

Berková, K., Králová, A. and Krejčová, K. (2017) 'Application of the metacognitive strategy in economic education in the Czech Republic', *Turkish Online Journal of Educational Technology*, Special Issue for INTE 2017, pp. 373-380.

Berková, K. and Krejčová, K. (2016) 'Effect of teachers´ abilities on students´ motivation with varying levels of intellectual abilities in the economics', *Journal on Efficiency and Responsibility in Education and Science,* Vol. 9, No. 3, pp. 81-87. https://doi.org/10.7160/eriesj.2016.090304

Boekaerts, M. (2004) Motivation to learn, Geneva: International Bureau of Education.

Brožová, H., Horáková, J. and Fiedler, J. (2018) 'Lecturers' managerial competencies important for students at the Czech university of life sciences', *Problems of Education in the 21st Century,* Vol. 76, No. 4, pp. 465-482.

Dytrtová, R. and Krhutová, M. (2009) Učitel. Příprava na profesi, Prague: Grada.

Findik-Coskuncay, D., Alkis, N. and Ozkan-Yildirim, S. (2018) 'A structural model for students' adoption of learning management systems: An empirical investigation in the higher education context', *Educational Technology & Society,* Vol. 21, No. 2, pp. 13-27.

Fontana, D. (2014) Psychologie ve školní praxi, Prague: Portál.

Kim, H. and Rehg, M. (2018) 'Faculty performance and morale in higher education: A systems approach', *Systems Research and Behavioral Science*, Vol. 35, No. 3, pp. 308-323. http://dx.doi.org/10.1002/sres.2495

Krejčová, K. and Berková, K. (2016) 'Abilities of Economic Subjects' Teachers and Their Impact on Students' Motivation', *Proceedings of the 13th International Conference Efficiency and Responsibility in Education 2016, Prague*, pp. 286-292.

Koucký, J. and Bartušek, J. (2011) Demografický vývoj a projekce výkonů vysokých škol, Prague: Univerzita Karlova v Praze.

Krpálková Krelová, K. (2014) 'Quality management systems in economic education and school enrollment rates as a parameter of quality school', *Integration of Economic Education, Prague*, pp. 106-111.

Lau, K.H., Lam, T., Kam, B.H., Nkhoma, M., Richardson, J. and Thomas, S. (2018) 'The role of textbook learning resources in e-learning: A taxonomic study', *Computers & Education,* Vol. 118, pp. 10-24. http://dx.doi.org/10.1016/j.compedu.2017.11.005

Mulač, P. and Mulačová, V. (2014) 'Výsledky dotazníkového šetření k řízení kvality českých neuniverzitních vysokých škol', *Proceedings of the 6th International Conference Competition, Jihlava*, pp. 165-173.

Ogrodzka-Mazur, E., Szafrańska, A., Malach, J. and Chmura, M. (2017) 'The Use of E-learning Resources by Academic Teachers – a Polish- Czech Comparative Study', *The New Educational Review,* Vol. 50, No. 4, pp. 169-185. http://dx.doi.org/10.15804/tner.2017.50.4.14

Pasiar, L., Berková, K., Krejčová, K., Pavera, L., Čonková, A. and Chmelárová, Z. (2015) Osobnost' učitel'a v ekonomickom vzdelávaní, Bratislava: EKONÓM.

Plháková, A. (2003) Učebnice obecné psychologie, Prague: Academia.

Savenkova, L.G. and Olesina, E.P. (2018) 'Strategic planning as a factor of improvement of contemporary education quality management', *Quality - Access to Success,* Vol. 19, No. 164, pp. 82-87.

Stranovská, E., Lalinská, M. and Boboňová, I. (2018) 'Teachers motivation towards assessment of their professional competences', *Problems of Education in the 21st Century,* Vol. 76, No. 4, pp. 561-574.

Šipikal, M. and Némethová, V. (2017) 'Performance based financing of universities – case study of Slovaika', *Scientific Papers of the University of Pardubice, Series D, Faculty of Economics and Administration*, Vol. 24, No. 41, pp. 169-180.

Tan, P., Wu, H., Li, P. and Xu, H. (2018) 'Teaching management system with applications of RFID and IoT technology', *Education Sciences,* Vol. 8, No. 1, pp. 1-13. http://dx.doi.org/10.3390/educsci8010026

Windham, D. M. (1988) 'Effectiveness indicators in the economic analysis of educational activities', International Journal of Educational Research, Vol. 12, No. 6, pp. 575-665. https://doi.org/10.1016/0883-0355(88)90011-0

Zygaitiene, B. and Buivydaite, E. (2017) 'Preparation of technology education teacher in Lithuania, Finland and Great Britain', 10th International Scintific Conference on Rula Environment, Education and Personality (REEP), Jelgava, pp. 408-415.

ON THE PROBLEM OF GENERATING A LARGE NUMBER OF COMPARABLE TEST VARIANTS

Mikuláš Gangur[1], Miroslav Plevný[2]✉

[1]Faculty of Economics, University of West Bohemia in Pilsen, Plzeň, Czech Republic

✉[2]Faculty of Economics, University of West Bohemia in Pilsen, Plzeň, Czech Republic, plevny@kem.zcu.cz

Highlights

- *The algorithm of generation a large number of content identical test variants in acceptable time is presented*
- *The tests may consist of more questions stemming from various thematic fields*
- *The content of individual test variants is identical but the similarity (congruence) is minimal*
- *The time of generation was shortened from hours to seconds in complex cases*

Abstract

The paper presents a possible way of solving the problem of creating more test variants for a large number of students divided into groups. The proposed solution may consist in introducing a parameterized automatic test generator. The principle of an automatic parameterized test generator is shown. The process of the question tree construction according to the increasing numbers of question in the banks of the particular subjects leads to a combinatorial explosion. This often results in excessive time of generation of the different variants of tests. To solve this problem, a heuristic method based on a pre-processing stage that precedes the construction of the searching tree is proposed. Further, the results of the experiments comparing the time of the test generation and the congruence of the test variants generated by the algorithm either using or non-using this heuristics are presented. According to these results the use of the generator with the proposed heuristics provides a considerably shorter time of generation, and the congruence of the generated test variants is even better in most cases.

Article type

Full research paper

Article history

Received: October 16, 2018

Received in revised form: December 4, 2018

Accepted: December 19, 2018

Available on-line: January 1, 2019

Keywords

automatic test generation, combinatorial explosion, evaluation methodologies, XSL transformation

Gangur M., Plevný M. (2018) "On the Problem of Generating a Large Number of Comparable Test Variants", *Journal on Efficiency and Responsibility in Education and Science, Vol. 11, No. 4*, pp. 78-84, online ISSN 1803-1617, printed ISSN 2336-2375, doi: 10.7160/eriesj.2018.110402.

Introduction

One of the most common forms of ascertaining (or measuring) the level of the gained knowledge is testing. Apart from the issue of the correct and well-balanced composition of the test, which we are not going to deal with herein, there is a specific issue of the need for designing a number of comparable variants for the same purpose of testing. A requirement calling for the creation of more variants of a test of the same type is quite common in this respect. This may be the case of testing the knowledge of a large number of students divided into groups. The same need arises when both the tests and the corrective tests are prepared simultaneously, or when the tests are repeated with a time delay. Another case is testing the knowledge in an e-learning environment where every student handles the test in a different time and it is therefore necessary to create a variant of the test for each student individually (Rosman and Buřita, 2014).

The problem of compilation of such a test set consists mainly in meeting the following requirements:

- All variants of the test should contain the same number of questions, as well as the same intensity expressed in points for each thematic field of the test;
- The similarity of the individual variants of the test should be minimal (a minimal congruence requirement).

Probably the biggest problem, when processing the necessary amount of the test variants manually (i.e. collection of tasks from each guarantee of an individual thematic field, the assembly of question groups for each individual variant of the test, etc.), consists in receiving the materials in different formats, in unbalanced difficulty of manually compiled variants of the test, and, apart from that, a time-consuming detection and correction of errors arising from editing the final form of the test. The use of MS Word templates for creating questions by the guarantees of thematic fields is seen as underperforming. Automatic generation of the test variants seems to be a possible solution to this problem.

Generating the necessary amount of variants of the given test may be facilitated by applying the computing technology. However, in connection with the above, it is necessary to solve a few problems. From the point of view of the generated tests mainly the following must be considered:

- guarantee of a comparable structure; and
- comparable level of difficulty.

The key issue of comparability of tests in terms of difficulty has been addressed in other publications. Authors Klůfa and Kaspříková (2012) use probability approach to solve the problem by binomial distribution and to answer the questions concerning the number of correct answers or the probability that the number of correct answers exceeds a given number. Similarly, Klůfa (2016) analyses the point number differences in the mathematics test among several variants of entrance examination test according to difficulty of variants. He studies how the results of entrance examinations depend on test variants.

For mentioned purpose the computing technology has been using for a long time. In the era of the Internet and the on-line technologies another significant advantage of using the computing technology for the purpose of testing appeared, namely the remote knowledge verification using on-line connection (Hürst, Jung and Welte, 2007; Niazi and Mahmoud, 2000).

The objective of this article is to present a proposal of the system for the automatic test generation according to the set requirements related to the structure (score, number of questions, issues tested) and the requirements concerning the mutual relations of the tests, i.e. the disparity of their contents. The algorithm of the test construction is described and a method for a solution of the problem of the combinatorial explosion during the question tree construction is proposed. The proposal of the above generator is based on a few input requirements for the basic characteristics and the generator functionality. These requirements are described in detail by (Gangur, 2014). The most important requirements are:

- the need for the existence of a simple control mechanism when setting parameters and generating tests;
- the possibility of the text contents structure parameterization, i.e. the possibility of determining the total score in the test and the number of questions stemming from the individual thematic fields;
- the possibility to ensure the minimal congruence of the test contents with regard to the random choice of questions from the question bank.

The above requirements were the starting points in the process of searching for such a system or, as the case may be, in applying the principles of the already designed and published systems.

In the next section of this paper the related work is presented, the functionality of the generator is briefly introduced, and the following parts describe the methods and algorithm of the generator implementation as depth/first tree searching with a backtracking mechanism. Then, our new proposal of a heuristic method for the solution of the combinatorial explosion is introduced. This explosion results from the question tree construction with respect to the number of questions. The use of the aforementioned generator remains only in theory without an application of the designed heuristics, due to the time-consuming calculation. Using this heuristics a gap between the theoretical and practical utilization of the generator is overcome, enabling the real use of the generator in practice. Finally the outcomes of experiments of the test generating process are presented while the results of the processes either using or non-using this heuristics are compared.

Materials and Methods

Related work

The issue of the automated test generating has been dealt with in a number of publications; see e.g. Brusilovsky and Pathak (2002), Sung, Lin and Chen (2007) or Zeng et al. (2013).

The similar difficulty of each variant is considered as a key issue of test variants design. Contributions dealing with this problem use various approaches to solve it. For example Klůfa and Kaspříková (2012) reflect the results of statistical analysis using probability for evaluation of appropriateness of test variants. Foltýnek (2009) applied another approach that enables to compare test variants results according to different difficulty level using scoring process and correctness coefficients.

Automated creation of adaptive tests with regard to the level of knowledge of the individual students is an independent field in which intensive research is being carried out (Mine, Shoudai and Suganuma, 2000; Kapusta, Munk and Turčáni, 2010). Fakhrusy and Widyani (2017) developed Moodle plugins to generate quiz as a part of LMS using genetic algorithm. Nuthong and Witosurapot (2017) focused on diverse difficulty of quizzes and proposed the 5-level difficulty ranking score using a hybrid similarity measurement approach to increase the number of usable generated quizzes and their sensible generation.

Seemingly simple issues of the automated test generation controlled by parameterized requirements concerning the test structure are not paid so much attention with regard to the quantity of publications on this topic. Authors Yang, Wu and Wang (2008) proposed and implemented a robust system with adaptive elements for administration and a follow up selection of the test questions from the database with regard to the previous test results is described. The system enables a random choice of the test questions with regard to the set parameters, such as the percentage of the required type of questions (e.g. multi-choice, open questions) or the fields (knowledge points) out of which questions are selected. In the key issue of the choice of questions the system uses a complicated mechanism of arithmetic calculations which ensures meeting the set requirements for the test structure. The system is extensive and with regard to the process of choosing the questions and feeding the question database it may seem difficult for the users.

None of the above systems deals with the problem of insertion and namely the typesetting of the mathematical text. The authors Tomas and Leal (2013) deal with the issues of the mathematical text by means of an external application. So, as to finalize the creation of the tests, the above authors use some functionalities of a web application for the presentation and evaluation of the mathematical expressions.

The above described systems meet the basic requirements for the test creation from the randomly chosen questions with regard to the set parameters of difficulty and coverage of various fields of issues to be studied. These are complex and extensive systems covering a number of other functionalities and requiring a time consuming creation of a question bank. In most cases, generators do not deal with the issue of the mathematical text typesetting and they are not quite flexible in the matter of the choice of the generated tests output format. The majority of the above stated tools as well as other examined instruments only generate online web tests.

The functionality of the generator

In this chapter we will discuss the functionality of the proposed automatic test generator along with methods and algorithms of the implementation of such a generator. Attention will be paid to the solution of the combinatorial explosion with focus on the mutual congruence of the generated tests.

Let us, first of all, describe the final product of the process of generation, which is the necessary amount of variants of the required test. The input data here are the ***source questions*** in the ***question bank***. The following attributes have to be entered for each question:

- ***thematic field*** – a thematic area related to the given question; for each thematic field the required number of questions and the overall number of points for this field must be entered as the input parameters of the process of generation;
- ***score*** - number of points awarded if the answer is correct, this value should express the difficulty of a question;
- ***group*** - it determines whether a question is or is not incorporated in the test in context with other questions – for more detail see following subchapter).

Resulting test consists of questions generated within the individual *thematic fields.*

The functionality of the generator is controlled by a set of input parameters. We used the following parameters as the basic input generator parameters which then determine the system functionality:

- the total number of questions stemming from the individual thematic fields;
- the total score stemming from the individual thematic fields;
- the format of the resulting tests;

- the total number of the generated tests;
- the number of the tests in a package to be used simultaneously.

By means of the first two parameters it is possible to select the quantity of questions stemming from the given field in the entire test and at the same time to select its difficulty by a suitable combination of the number of questions and the score for the given thematic field. The generator supports the distributed creation of the individual fields of questions by various creators who can save the final question bank in an online repository of questions used by the generator. The administrator then controls the final tests generation. This approach enables, in some special cases (entrance tests and such like), hiding the contents of the complete test from the individual creators and letting only one authorized person create the test.

The strength of the generator consists in the possibility of selecting a template for generating the required output format. The test itself is generated in the proposed universal XML format, and by means of the XSLT processor it is, with the help of the inserted transformation template, transformed into the required output format (Kosek, 2013). The possibility of selecting this output format is flexible and it enables the user to create his/her own template and to generate his/her own output format (LaTeX, AcroTeX, Moodle XML).

The parameters determining the total number of the generated tests and the number of the tests in a package to be used simultaneously also control, among other things, the format of other generator outputs, namely the calculation of the mutual percentage congruence of the test variants, and the suggestion of the most suitable combinations of the test variants to be used simultaneously. The administrator, with the help of these suggestions, tries to compile the tests so that there are tests with the lowest level of congruence of questions between the individual rounds. The test questions are selected randomly and some questions, with regard to the required total number of questions in comparison with the number in the bank of a given field, may be repeated in the tests.

Even the question banks stemming from the individual fields can be listed in the generator outputs. The possibility of simple creation of such a question bank by means of freely available editors is one of the requirements for the generator functionality.

The control information related to the individual questions can be seen as another parameter influencing the test compilation. It determines, apart from the evaluation of a question by scoring, also listing a question in a group of questions. The group of questions enables similar questions not to be listed in one test and, at the same time, to list more questions with the joint settings in one group.

Another functionality of the generator considers congruence among test variants. As support for the prevention of undesirable cooperation among the examinees the generator considers the percentage congruence of tests and proposes combinations of the individual test variants to be grouped together. The percentage congruence of two variants is defined as a ratio of the number of identical questions in the considered variants and the total number of questions.

The proposal for the composition of variants in the individual rounds, i.e. the test packages, results from the requirement for the minimal congruence between the individual rounds. This limits the possibility of influencing the test as a result of possible communication of the examinees in the time gap between rounds when the examinees from one round may pass on as little information related to the particular questions as possible to the examinees in the following round.

The question structure and information control

The current version of the presented generator uses the Aiken question format (Aiken, 2013) and it can be extended by the possibility of the questions with a short or numerical answer and by the possibility of inserting more correct answers in case of the multiple response questions.

Each question is introduced by a tag with an abbreviation of the thematic field to which the question belongs (see ‘OV’ in the listing below). The tag contains control information influencing the listing of questions in the compiled test. Within this information the question bank creator determines the evaluation of the question by score and the group to which the creator lists the question.

```
    <OV score="2" group="381"> The
objective function for achieving the highest
total possible number of the manufactured
products in the linear mathematical model
of an optimization task for the above stated
settings can have the following from:
    A) <math>\max z = \sum\limits_{i=1}^{n}
w_i</math>
    B) <math>\max z = c_j \sum\limits_
{i=1}^{n} w_i</math>
    C) <math>\max z = \sum\limits_{i=1}^{n}
c_{ij} w_{ij}</math>
    D) <math>\max z = \sum\limits_{i=1}^{n}
\sum\limits_{k=1}^ {p} b_{jk} w_j</math>
    E) <math>\max z = \sum\limits_{i=1}^{n}
c_j w_j</math>
    ANSWER: A
```

The numerical code identifying the group is important. The digits of this code control the listing of the question in the stage of constructing the test according to the following scheme:

- Group 0 - the question can be listed without limitation;
- Group <1 – 99> - questions with the same number are not listed together in one test;
- Group <100 – 999> - group questions; mostly it is more questions with joint settings;
    - questions with the same first digit and different second digit are not listed in the same test;
    - questions with the same first and second digit belong to the same group and either all of them are listed in the test or none of them at all;
    - the last digit determines the order of questions in the group; the first one is often a question with the joint settings.

One of the other features of the generator is the possibility to insert the mathematical text into the text of a question or, as the case may be, also the exact listing (e.g. algorithm listing and such like) as well as a figure in the JPEG format as a complement to the task settings (Gangur, 2011; Gladavská and Plevný, 2014).

The algorithm of the test assembling

In case we require the compilation of the test out of the questions based on the set criteria and with regard to the question control information (group) the algorithm of the recursive depth-first search of tree is applied. It is the so called backtracking algorithm which selects, out of the questions for the given field, one or more questions (according to the group number) and it always checks whether the criteria of the total number of questions and the required total score are met. If one of these parameters is exceeded, it recurs by one question (more questions) and selects another one.

The core of this algorithm is the ***combination*** recursive function. The input in this function, when it is called for the first time, is the list of the question bank suitable for the given field. Random permutation and question selection is applied in case of this list and therefore the order of questions and the depth-first search of tree are always different. The following listing shows the headline of the applied function and its first call.

```
    function combination($list, $current_
list, $num_points, $deep)
    new_test_list = combination($question_
bank, Array(), 0, 0)
```

In case of further recursive calls of the function this list is entered in the function without the questions that had already been used. In this sense, the current list is also an input parameter in which selected questions are stored (in case of the first call the list is empty). Other parameters are: the score of the questions currently inserted into the test and the depth of tree which represents the number of questions in the compiled test. Upon the first call both the values are null. With each question (group of questions) being added the depth of tree gets higher.

Out of the list of free questions the recursive function call creates the rest of the list of the tested questions. The recursion ends upon achieving the required score and the number of questions for the given field. In case of the retrospective finishing of the individual calls of the function a list of the test questions is formed starting from the back and moving forward and at each level this tail of the list is added to the currently selected question or the question group and like this a new tail of the list is created at the given call level.

If, upon the function call, the values of score or number of questions are exceeded, the selected question (group of questions) is not accepted and another question in the list of free questions is chosen, until the list is empty. After that the algorithm recurs back by one level of the call (backtracking), and it selects another question out of the list of the free questions at the given level.

By means of the above described procedure of backtracking the depth-first search of tree is implemented. Upon returning back to the first call level the whole list of questions according to the set criteria is created if the finishing condition is met.

The process of depth/first searching algorithm according to the number of used questions leads to the combinatorial explosion and it takes too much time. When increasing the number of the source questions for a thematic field as well as the number of the demanded questions for this field the time demand of the generation process increases significantly. In many cases, this time is expressed in the order of tens of minutes (or hours sometimes). This can be very annoying for users, and therefore it is necessary to solve this problem. A proposal of a possible way to solve this problem is described in the next part.

The solution of the problem of combinatorial explosion

One of the possibilities to solve the problem of the combinatorial explosion of questions tree is to decrease the whole number of questions in one thematic field. This approach narrows the selection of different questions and increases the possibility of higher congruence among the generated tests.

We, on the other hand, propose the solution that also decreases the number of questions put in the process of depth-first tree searching, but the algorithm randomly selects these questions from the bank of all questions of one thematic field. If the assembling of the questions list for the generated test according to the input parameters setting is not successful, the process continues with the selection of new questions from the rest of questions in the bank. When the bank of questions for the considered field is emptied all questions are returned to the bank and the process starts again with the new selection.

In the preselection we take into account the exclusive questions with respect to the group number (1 – 99), and at the same time the preselection process controls the whole number of the demanded points for the field. This number of points is satisfied by the preselection of questions individually for sets of different *n*-points questions. For example 10 points can be assembled from four 3-points questions, five 2-points questions, and ten 1-point questions. In the same way the question list is constructed for every new test variant, and it contributes to the low congruence among different test variants.

The step of the random selection is implemented as the selection of all the remaining questions in the bank in the permuted order. This preselection has to reflect the group questions, i.e. if the one question of the group is selected the other ones have to be included to the selection. The described process has quadratic time complexity according to the number of the preselected questions and the number of all the source questions.

The algorithm of the preselection is described in the next steps. The input to this preselection process is a permuted list of the source questions.

```
Step 1. Take the input current list of the
        source questions as pl.
Step 2. Create an array of maximal
        numbers of questions for every
        n-point question according to the
        prescription floor (demanded points
        for field / n)+1.
Step 3. While selected number of n-point
        questions is not greater or equal
        to the maximal numbers of questions
        determined in step 2 for every n or
        all questions of list pl are checked
        do
        i.  Take question from source list
            pl and fill n as a score of
            question.
        ii. If (number of n-question is
            less than the maximal number
            of n-questions) and group of
            question < 100 is not in the
            resulting question list rl
            then
            a. if the group of question >=
               100 then select all questions
               of the group and add them to
               question;
            b. add question or all selected
               question in group to the
               resulting question list rl.
            c. remove the selected
               question(s) from the list of
               the source questions pl.
Step 4. Return the resulting list of
        questions rl.
```

The output from the preselection is the list of questions and it is one of the input parameters for the *combination* function (see subchapter *The question structure and information control*). If the assembling of questions for the field is successful the selected questions are removed from the permuted list of the source questions and for assembling the next variant the preselection process selects questions from the modified list of the source questions. If the assembling is not successful the list of the source questions is fulfilled with all the questions and the preselection is applied on this list. If the assembling is not successful on this new preselected list the generator stops without result.

The proposed heuristics not only decreases the time consumption but also, with respect to selected questions pools (that are mostly mutually exclusive), supports lower congruence among particular variants of the generated tests. The comparison of the results of the generation processes using or not using the proposed heuristics is presented in the next part.

Results of experiments

The outputs of the generating process in different output formats according to the input parameters are presented by (Gangur, 2014). The current version of the generator offers XSL templates mainly for the formats Moodle XML, LaTeX and AcroTeX (interactive PDF). In case of TeX format (LaTeX and AcroTeX) it is possible to create PDF documents by means of the post-processor methods.

Next the computing times and the congruence among generated variants are presented. The generator was used for the same source of questions with the same input parameters, i.e. the same number of the demanded questions for every field and the same demanded score for every field. In the first case the generator was implemented without the above described heuristics, and in the second case with this heuristics. The data in the table 1 show the numbers of question in the banks of the particular subjects (thematic fields) together with the values of the input parameters.

Particular subject (thematic field)	Bank size (No of source questions)	No of demanded questions	Demanded score of questions
Economics	176	10	20
Business economy	140	11	20
Management	81	5	10
Marketing	106	6	10
Business finance	70	5	10
Accounting	70	5	10
Management science	59	4	10
Statistics	69	5	10
Average	96.375	6.375	12.5

Table 1: Numbers of questions in banks and input parameters values, 2017 (source: own calculation)

The output coefficients of comparison are the congruence of variants and the speed (time complexity) of the generating process. In one experiment 10 or 5 test rounds are processed. The outputs of the experiment are the average time per one round (AT) for each compared algorithm (without heuristics - NH, and with heuristics - H), and the average coefficient of congruence among particular variants per one round (AC).

The input parameters for experiments are as follows:

- The number of test rounds.
- The number of the source questions represented as an average number of the source question per one thematic field (ANSQ). This number is determined by the selection of the configured percent part of the basic source files (see example of questions number in the table 1). More source questions generally imply larger computing time and the decrease of the resulting congruence among particular variants of test.
- The number of the demanded questions for every thematic field is represented by the average number of the demanded question per one field (ANDQ). These numbers are configured by the absolute number of questions for every field. As with the previous parameter ANSQ the higher number of the source questions generally imply the larger computing time and the decrease of the resulting congruence among particular variants of test.
- The number of the demanded variants of the test is configured as the absolute number of variants. In the experiments this number was set to 6 variants (see later).

Version	Algorithm with heuristics (H)						Algorithm without heuristics (NH)					
	A	B	C	D	E	F	A	B	C	D	E	F
A	-	0.00	0.00	0.00	0.00	0.00	-	5.88	7.84	7.84	1.96	5.88
B	0.00	-	0.00	0.00	0.00	0.00	5.88	-	0.00	3.92	3.92	1.96
C	0.00	0.00	-	0.00	0.00	0.00	7.84	0.00	-	3.92	1.96	9.80
D	0.00	0.00	0.00	-	0.00	0.00	7.84	3.92	3.92	-	7.84	5.88
E	0.00	0.00	0.00	0.00	-	0.00	1.96	3.92	1.96	7.84	-	13.73
F	0.00	0.00	0.00	0.00	0.00	-	5.88	1.96	9.80	5.88	13.73	-

Table 2: Values of mutual congruence using algorithm H and NH [%], 2017 (source: own calculation)

The values of both parts of the table 2 are summarized and represented by one coefficient of the average congruence. The lower value of the congruence means a lower level of mutual similarity of variants. In our case the values of this coefficient are 0.00% for the generation processes using the proposed heuristics (H), resp. 5.49% for the generation processes non-using the proposed heuristics (NH).

The outputs are the average congruence of one test round of the generator and the time of generating the demanded number of variants. The searching tree is constructed for every thematic field of every variant and that's why the complexity increases linearly according to the increasing number of the demanded variants. With respect to the number of the source questions the average congruence among variants increases when the demanded variants increase. If we process more than one generator round in one experiment it is possible to characterize the experiment results as the whole average congruence per one round and the whole average time per one round. The example of such an experiment is illustrated by the following table 3 where 100% of all questions were used, i.e. on average 96.375 questions per one field and on average 6.375 demanded questions per one field (see table 1). In table 3 the average congruence among variants in % for every round and both algorithms are stated in the first two rows. The computing time of generating the demanded variants in seconds for every round and both algorithms are presented in the last two rows. The last column shows the whole averages per one round.

	N	1	2	3	4	5	6	7	8	9	10	Avg.
Congru-ence [%]	H	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	NH	5.49	6.14	7.04	7.44	7.97	6.79	7.71	6.53	6.01	8.49	7.01
Time [s]	H	0.21	0.19	0.23	0.26	0.25	0.24	0.30	0.37	0.66	1.19	0.39
	NH	0.76	0.73	1.27	0.92	1.00	7.48	2.09	2.00	3.67	15.38	3.53

Table 3: Average congruence and the time of generation, 2017 (source: own calculation)

The described experiment can be characterized by the following values of the input and output parameters:

ANSQ= 96.375

ANDQ= 6.375

AC = 7.01% for NH, 0% for H

AT = 3.53 [s] for NH, 0.39 [s] for H

In the implemented experiments the values AC (table 4) and AT (table 5) were measured for different values of ANSQ from 100% to 20% of the basic file of questions and for the increasing values of ANDQ.

For every ANSQ and ANDQ the pair NH | H of values for the algorithm without heuristics and algorithm with heuristics is presented. In every experiment 10 rounds for ANDQ=6.375 and 5 resp. 3 rounds for ANDQ=11.25 resp. ANDQ=20 were processed with respect to the time complexity of computing. If the test was not able to assemble according to the demanded criteria (the number of points and the number of questions for every thematic field) in the given round, the random selection form of the source questions was implemented again until the test was assembled or the experiment finished after the configured time (it was set to 2 hours in all the experiments provided).

ANSQ	ANDQ = 6.375		ANDQ = 11.25		ANDQ = 20.00	
	NH	H	NH	H	NH	H
96.375 (100%)	7.15	0.00	Inf	4.83	Inf	21.54
86.50 (90%)	7.97	0.00	23.53	11.16	Inf	25.60
76.75 (80%)	8.34	0.14	19.04	13.35	Inf	35.98
67.25 (70%)	10.57	0.80	21.91	16.31	Inf	x
57.50 (60%)	11.88	1.96	24.88	20.18	Inf	x
48.00 (50%)	15.20	4.30	31.14	29.36	Inf	x
38.25 (40%)	18.20	6.54	38.74	39.33	x	x
28.50 (30%)	25.45	17.41	x	x	x	x
19.00 (20%)	37.38	38.17	x	x	x	x

Marking used in the table 4 and 5:

x - *the test was not assembled with respect to the demanded criteria according to a small number of the source questions,*

Inf - *the computing finished according to overtime (2 hours).*

Table 4: Average congruence (AC) without heuristics (NH) and with heuristics (H) [%], 2017 (source: own calculation)

ANSQ	ANDQ = 6.375		ANDQ = 11.25		ANDQ = 20.00	
	NH	H	NH	H	NH	H
96.375 (100%)	1.03	0.16	Inf	0.30	Inf	187.13
86.50 (90%)	1.43	0.16	115.34	0.26	Inf	92.37
76.75 (80%)	1.01	0.15	111.95	0.23	Inf	5.37
67.25 (70%)	0.65	0.13	7.41	0.27	Inf	x
57.50 (60%)	0.83	0.12	7.28	0.25	Inf	x
48.00 (50%)	0.41	0.13	10.70	0.20	Inf	x
38.25 (40%)	0.27	0.10	0.72	0.27	x	x
28.50 (30%)	0.20	0.12	x	x	x	x
19.00 (20%)	0.13	0.10	x	x	x	x

Table 5: Average time (AT) without heuristics (NH) and with heuristics (H) [s], 2017 (source: own calculation)

Discussion

As it is apparent from the values in the tables 4 and 5 for the increasing number of the demanded questions (ANDQ > 6.375) and the decreasing number of the source questions the assembling of the test was not successful in the configured max time of the process. The presented values depend on the structure of the source questions in the thematic field according to the enlistment of questions to the groups and the mutual excluding in one test as well as the number of the tied questions and the tied groups. That's why it is suitable to compare the obtained values for NH algorithm and H algorithm that were computed in relation to the same data, i.e. to the same structure of the source questions.

Despite the previous statement it is obvious that with increasing ANDQ the congruence increases as well as the time complexity. Especially the time complexity can be characterized as $O(K(\text{ANSQ}, \text{ANDQ}))$, where K is the binomial coefficient, i.e. the time complexity depends above all on the number of the source questions (ANSQ) and the number of the demanded questions (ANDQ). The proposed heuristics decreases the number of the source questions in the process of the searching tree construction and this way the time complexity of the generation process decreases, but it does not increase the congruence. By contrast, in most comparisons the final congruence of the test generation the algorithm with heuristics is lower than the algorithm without heuristics.

The advantage of the proposed and in practice applicable generator is a simple system of parameter setting that controls generation in terms of difficulty and number of questions from each of the areas considered. Another advantage is the simple format of source texts for individual questions and thus a simple process of creating questions without the need to use a more complex SW. Optionally, it is possible to use already prepared questions that may be in plain text format. Among the undeniable advantages over similar systems is the choice of different output test formats. Tests can be generated in pdf format suitable for a written test printing, in html format for placing on the web for online testing or into the LaTeX structure for further processing. As limitations of the proposed solution, the impossibility to value the individual questions in view of the difficulty of other tasks in the final test may be considered. Each question is rated by a specific point value within a given area. On one side, the comparative difficulty level of each variant of the test is determined by the described approach of assembling the tests according to the given number of questions and the total number of score points representing difficulty from the given thematic field. On the other side, this assumes a correct assessment of the difficulty of questions on a scale of 1-3 points by the authors of questions from each individual field. The overall difficulty of variants of the entire test can be compared only with regard to this evaluation of individual questions by the good fit tests. Equally, the different level of students' knowledge is not considered by these tests (Klůfa, 2016). Compared to Klůfa and Kaspříková (2012), the results of previous tests are not taken into account by the proposed method, and the structure of the test is not modelled using estimated probability distributions.

Similarly, the possibility of more sensitive control of the difficulty of individual tasks is limited. All questions are of the MCQ type and are designed by the designers of the questions as static from the domain. Therefore, the system does not allow to variate the distractors (bad answers) dynamically with respect to adjusting the difficulty by choosing ontologically close distractors. The maximal number of questions in the area is limited also with regard to the system of grouping questions, which controls the placement of questions in the test within the required context. However, the proposed algorithm for classifying questions into groups is versatile, allowing the extension of the grouping interval used and thus increasing the maximal number of questions for the given area.

In the future, we consider extending the functionality of the test generator by the implementation the option into the generator to use not only statically specified but also parameterized questions.

Furthermore, we will focus on the issue of comparable difficulty level of individual variants both from the point of view of individual students' results and from the point of view of measurement and comparison of the difficulty of individual variants among themselves. This would extend the use of the generator significantly, and partially eliminate some of the above limitations.

Conclusions

The system described in this paper is a useful aid, namely in the process of preparing tests stemming from various fields, for example from the sphere of the entrance examination procedure applied in the authors' institution. When using a larger number of sources or demanded questions the original algorithm non-using the heuristics described above does not allow for assembling the demanded test variants in an acceptable time. On the contrary, the proposed heuristics allows the use of the generator in the real time.

The paper describes the functionality and the process of implementation of the generator of questions stemming from the set fields by means of a combination of questions according to the set criteria, i.e. namely the number of questions in the test for the given thematic field and the overall scoring of questions in the given field.

The solution of the generation process is based on the construction of a searching tree and its depth-first searching combined with the backtracking algorithm. The main input parameters are the number of source questions and the number of the demanded questions. These parameters most influence not only one of the most important final outputs of the generation process, i.e. the mutual congruence among particular variants of the test, but also the time complexity of the generation process. The depth-searching of the tree results in the combinatorial explosion and high increasing of time complexity. In some configurations of the input parameters, especially the source questions and the demanded questions, the generation process does not finish within the required maximal time limit of 2 hours. That is the reason why to propose the pre-process of preselection of a limited smaller number of questions without placing them back to the source questions. This preselection process also respects some input criteria, and it is able to satisfy them in the quadratic time complexity. If the generating process is unsuccessful because of the empty list of the source questions or a small number of questions in the list, the generator starts the process again with a new selection from all the newly permuted source questions. The proposed improving of algorithms decreases the complexity of computing namely with respect to the increasing number of the demanded questions, and contributes to the low congruence among different test variants. The mutual congruence of variants generated by the algorithm using the proposed heuristics is at least as good (and mostly even better) as the congruence of the test variants generated by means of an algorithm non-using this heuristics.

Acknowledgment

This work was supported by the University of West Bohemia in Pilsen under Grant No. SGS-2018-042.

References

Aiken (2013) *MoodleDocs: Aiken format.* [Online], Available: http://docs.moodle.org/22/en/Aiken_format [26 Jul 2016].

Brusilovsky, P. and Pathak, S. (2002) 'Assessing Student Programming Knowledge with Web-based Dynamic Parameterized Quizzes'. A*ssociation for the Advancement of Computing in Education (AACE)*, pp. 1548-1553.

Fakhrusy, M.R. and Widyani, Y. (2017) 'Moodle plugins for quiz generation using genetic algorithm'. *2017 International Conference on Data and Software Engineering (ICoDSE)* Palembang, Indonesia, INSPEC Accession Number: 17578381. http://dx.doi.org/10.1109/ICODSE.2017.8285882

Foltýnek, T. (2009) 'A New Approach to the Achievement Test Items Evaluation: the Correctness Coefficients'. *Journal on Efficiency and Responsibility in Education and Science*, vol. 2, no. 1, pp. 28-40.

Gangur, M. (2011) 'Automatic generation of cloze questions'. C*SEDU 2011 - Proceedings of the 3rd International Conference on Computer Supported Education*, Portugal, SciTePress - Science and Technology Publications, pp. 264-269.

Gangur, M. (2014) 'Automatic Parameterized Generation of Test'. *DIVAI 2014: 10th International Scientific Conference on Distance Learning in Applied Informatics*. Prague, Wolters Kluwer, pp. 55-64.

Gladavská, L. and Plevný, M. (2014) 'Problems of automatic generation of questions for the purpose of testing the knowledge in a management science course'. *DIVAI 2014: 10th International Scientific Conference on Distance Learning in Applied Informatics.* Prague, Wolters Kluwer, pp. 325-335.

Hürst, W., Jung, S. and Welte, M. (2007) 'Effective learn-quiz generation for handheld devices'. Pro*ceedings of the 9th international conference on Human computer interaction with mobile devices and services*, ACM, pp. 364-366.

Kapusta, J., Munk, M. and Turčáni, M. (2010) 'Evaluation of adaptive techniques dependent on educational content'. *4th International Conference on Application of Information and Communication Technologies, AICT2010*, INSPEC Accession Number: 5611791. http://dx.doi.org/10.1109/ICAICT.2010.5611791

Klůfa, J. and Kaspříková, N. (2012) 'Multiple Choice Question Tests for Entrance Examinations - A Probabilistic Approach'. *Journal on Efficiency and Responsibility in Education and Science*, vol. 5, no. 4, pp. 195-202. http://dx.doi.org/10.7160/eriesj.2012.050402

Klůfa J. (2016) 'Comparison of the Test Variants in Entrance Examinations'. *Journal on Efficiency and Responsibility in Education and Science,* vol. 9, no. 4, pp. 111-116, http://dx.doi.org/10.7160/eriesj.2016.090404

Kosek, J. (2013) *XML for everyone.* [Online], Available: http://www.kosek.cz/xml/index.html [20 May 2016].

Mine, T., Shoudai, T. and Suganuma, A. (2000) 'Automatic Exercise Generator with Tagged Documents Considering Learner's Performance'. *Association for the Advancement of Computing in Education (AACE)*, Chesapeake, VA, pp. 779-780.

Niazi, R. and Mahmoud, Q.H. (2000) 'A Web-based Tool for Generating Quizzes for Mobile Devices'. *A poster at the 39th ACM Technical Symposium on Computer Science Education (SIGCSE).* Portland, OR.

Nuthong, S. and Witosurapot, S. (2017) 'Enabling fine granularity of difficulty ranking measure for automatic quiz generation'. *9th International Conference on Information Technology and Electrical Engineering (ICITEE)*, Phuket, Thailand, pp. 1-6. http://dx.doi.org/10.1109/ICITEED.2017.8250492

Rosman, P. and Buřita, L. (2014) 'Concept of the computer science course and some aspects of ICT integration into education'. *E&M Ekonomie a management*, vol. 17, no. 3, pp. 169-180. http://dx.doi.org/10.15240/tul/001/2014-3-013

Sung, L.-C., Lin, Y.-C. and Chen, M.C. (2007) 'An Automatic Quiz Generation System for English Text'. *7th IEEE International Conference on Advanced Learning Technologies, Proceedings*, Niigata, Japan, INSPEC Accession Number: 9868711. http://dx.doi.org/10.1109/ICALT.2007.56

Tomas, A.P. and Leal, J.P. (2013) 'Automatic generation and delivery of multiple-choice math quizzes'. P*rinciples and Practice of Constraint Programming. CP 2013. Lecture Notes in Computer Science,* Vol. 8124. Springer, Berlin, Heidelberg, pp. 848-863. http://dx.doi.org/10.1007/978-3-642-40627-0_62

Yang, A., Wu, J. and Wang, L. (2008) 'Research and design of test question database management system based on the three-tier structure'. *WSEAS Transactions on Systems*, vol. 7, no. 12, pp. 1473-1483.

Zeng, J., Sakai, T., Yin, C., Suzuki, T. and Hirokawa, S. (2013). 'Automatic generation of tourism quiz using blogs'. *Artificial Life and Robotics*, vol. 17, no. 3-4, pp. 412-416. http://dx.doi.org/10.1007/s10015-012-0076-7

IMPACT OF ORDER OF DATA IN WORD PROBLEMS ON DIVISION OF A WHOLE INTO UNEQUAL PARTS

Jarmila Novotná[1✉], Martin Chvál[2]

[✉1]Faculty of Education, Charles University, Prague, Czech Republic, jarmila.novotna@pedf.cuni.cz

[2]Faculty of Education, Charles University, Prague, Czech Republic

Highlights

- *The presented results are of interest for mathematics education as well as for construction of tests*
- *Item response theory was used for evaluating differences in the difficulty of problems*
- *Analyses of pupils' solutions allow to get deeper insight into pupils' difficulties when solving word problems*

Abstract

The paper investigates how the order of numerical data in word problems on division of a whole into unequal parts affects achievement and reasoning of 14-16-year old pupils. The variable was altered in two word problems, in one of which also the context was changed (psychological variable) and in the latter "if-clause" is or is not used (linguistic variable). 182 pupils were involved in the experiment. The solutions were analysed quantitatively using Item Response Theory as well as qualitatively. The data suggest that pupils' success is affected by the order of numerical data in the statement in an unfamiliar context. The presence of "if-clause" in the statement was studied in a two-level problem. The order of numerical data played its role in case of formulations without "if". The results of the experiment are of interest for mathematics education as well as for construction of tests. The paper is an extended version of the paper by Novotná (2018).

Article type

Full research paper

Article history

Received: October 26, 2018

Received in revised form: December 10, 2018

Accepted: December 22, 2018

Available on-line: January 1, 2019

Keywords

Context, "if-clause" formulation, mathematics, order of numerical data, word problems

Novotná J., Chvál M. (2018) "Impact of order of data in word problems on division of a whole into unequal parts", *Journal on Efficiency and Responsibility in Education and Science, Vol. 11, No. 4*, pp. 85-92, online ISSN 1803-1617, printed ISSN 2336-2375, doi: 10.7160/eriesj.2018.110403.

Introduction

The area of word problems has been identified as an area of concern in many countries because of pupils' difficulties while solving them. Teachers and researchers report that word problems are one of the areas in which pupils show the poorest performance, see e.g. (Hembree,1992; Zohar and Gershikov, 2008). Wijaya, van den Heuvel-Panhuizen and Doorman (2015) give recommendations on how to improve the opportunities-to-learn while solving word problems as well as on how to conduct further research on this topic.

Novotná and Vondrová (2017) investigated the impact of context on Grade 8 pupils' choice of strategies when solving two missing value word problems with the same mathematical model but different contexts. Vondrová, Novotná and Havlíčková (2019) studied the impact of the order of numerical data, context, position of the unknown transformation and the length of the text in an additive word problem on the performance and reasoning of primary pupils from grades 4 and 5.

The study described here analyses only one type of word problems (problems on division of a whole into unequal parts) in terms of the impact of the variable "the order of numerical data in the word problem statement". The study is a part of a wider research within the project of the Grant Agency of the Czech Republic aimed at investigating variables influencing the difficulty of word problems. The project is unique at least in the Czech context because the research is conducted in cooperation of researchers in linguistic, mathematical education and psychology. This enables the project team to investigate the problem from different standpoints.

In this paper, the result of coding of word problem statement is called the *legend* (Novotná, 2010). The order of numerical data in the problem statement is called *proper* if the data appear in the sequence needed to solve the problem, otherwise it is called *mixed* (Vondrová, Novotná and Havlíčková, 2019). For the purposes of this paper, the *order of information recorded in a legend* is called *direct* if the data appear in the order used in the problem statement, otherwise it is called *indirect*. It is obvious that a proper order of numerical data does not imply that the solver records it in the proper legend and vice versa.

Searle, Lorton and Suppes (1974) reported that the order of data was a major predictive variable for pupils from Grades 4, 5 and 6 in problem solving. Similarly, Hembree's (1992) pointed better performance in case of proper order of data. Neither reference specifies how exactly the order of data was changed, though. On the other hand, Nesher, Hershkovitz and Novotná (2003) found that the effect of the order of presenting comparison relations was negligible in simple comparison problems.

Vicente, Orrantia and Verschaffel (2008) altered both mathematical and situational variables in word problems. The mathematically difficult problems included a combination of two inconsistent change situations; in the situationally difficult problems, the information about the initial moment of the action string was given at the end of the text. They found statistically significant differences in the performance of pupils in Grades 3 to 5 for both variables with a larger effect size coming from the mathematical difficulty level.

Soneira, González-Calero and Arnau (2018) analyse the solvers' translation of problem statements from natural language into language of algebra. They report that "the student performs a translation from natural language into algebraic language keeping the same order of the symbols in the equation as the key words in the statement, regardless of the meaning of the expressions" (2018: 44).

The paper answers two research questions: How does the order of numerical data in the problem statement in problems on division of a whole into three unequal parts influence the achievement and reasoning of Grade 9 (age 14-16) pupils? Do the solving strategies used by these pupils vary according to the numerical order of data in the statement? Attention in the analyses is also be paid to the impact of change in context or wording from the linguistic perspective on the choice of solving strategies.

Materials and Methods

Participants

As Novotná and Vondrová (2017: 280) report, pupils participating in the research are from four primary schools, which were purposefully sampled within the GAČR project (focusing on investigation of parameters influencing the difficulty of word problems). The schools were selected on the basis of Reports by the Czech School Inspection and of websites of the schools in with the aim of selecting schools with no specialisation, of medium size, attended by children from their immediate surroundings with a varied socio-economic background, with the percentage of children/foreigners not exceeding the average for the whole Czech Republic, not founded for children with special needs and placed in the area of outer Prague. An important criterion for inclusion of the school in the GAČR project was that the whole school would get involved. The *sample* used in the part of research presented here consists of 182 Grade 9 pupils. No selection of pupils was made; all the classes from the four schools of the same grade participated.

Materials

The pupils were given initial tests in mathematics and Czech language that enabled division of classes into equally abled groups. These groups were later given different versions of the same word problem differing in the order of numerical data in the problem statement or in the context or the presence of "if-clause". For the purpose of this text and based on our research questions, two word problems (labelled 9A and 9C in the following text) were selected. Both work with division of a whole into unequal parts where the whole and the relationships between the parts are given. In both problems, the variants are based on the same mathematical model. In Table 1 and 2, the statements are presented. In both, variants 1 and 3 have the same order of numerical data and variants 2 and 4 are their modifications in another context (9A1 versus 9A3 and 9A2 versus 9A4) or with another wording (9C1 versus 9C3 and 9C2 versus 9C4).

Familiar context	Unfamiliar context
9A1 Gymnastics, basketball and swimming clubs take place at the same time on Wednesday evening. Basketball club is attended by three times more children than gymnastics, swimming is attended by 114 more children than basketball. In total the three clubs are attended by 380 children. How many children are enrolled in each of the clubs?	**9A3** There are three political clubs in the parliament: Liberal, Conservative and Green. Each member of parliament can be member of only one of the clubs. The Liberal Club has three times more members than the Conservative Club, the Green Club has 114 more members than the Liberal Club. In total there are 380 members in all the three clubs. How many members are there in each of the clubs?
9A2 380 children are enrolled in three clubs: gymnastics, basketball and swimming. All these clubs take place at the same time on Wednesday evening. Basketball club is attended by three times more children than gymnastics, swimming is attended by 114 more children than basketball. How many children are enrolled in each of the clubs?	**9A4** 380 members of a parliament in a foreign country is divided into three political clubs: Liberal, Conservative and Green. Each member of parliament can be member of only one of the clubs. The Liberal Club has three times more members than the Conservative Club, the Green Club has 114 more members than the Liberal Club. How many members are there in each of the political clubs?

Table 1: Word problem 9A for Grade 9 in four variants (change of the order of numerical data in the statement and change of the context)

Formulated with "if" (linguistic variable)	Formulated without "if"
9C1 Students Kamil, Eva and David spent part of their holiday making money as tour guides at the Děčín palace. They earned 8 800 CZK, which they divided with respect to the number of tours. For one tour they earned 40 CZK. How did they divided the money if Kamil had 12 fewer tours than Eva and David had twice as many tours as Kamil?	**9C3** Students Kamil, Eva and David spent part of their holiday making money as tour guides at a palace. They earned 8 800 CZK. They divided this money with respect to the number of tours. For one tour they earned 40 CZK. Kamil had 12 fewer tours than Eva and David had twice se many tours as Kamil. How did they divide the money?
9C2 Students Kamil, Eva and David spent part of their holiday making money as tour guides at a palace. For one tour they earned 40 CZK. They divided the earned money with respect to the number of tours. How did they divide the money if Kamil had 12 fewer tours than Eva and David had twice as many tours as Kamil and in total they earned 8 800 CZK?	**9C4** Students Kamil, Eva and David spent part of their holiday making money as tour guides at a palace. For one tour they earned 40 CZK. They divided the earned money with respect to the number of tours. Kamil had 12 fewer tours than Eva and David had twice as many tours as Kamil. In total they earned 8 800 CZK. How did they divide the money?

Table 2: Word problem 9C for Grade 9 in four variants (change of the order of numerical data in the statement and change of formulation of the assigned conditions)

Note: The order of numerical data in the assignment, or context, or "if" formulation represents one of the mathematical, or psychological, or linguistic variables investigated in the GAČR project.

A brief analysis *a priori* in terms of solving strategies for both problems is presented below. Novotná et al. (2013) give a list of heuristic solving strategies used for this analysis. Both "school" algebraic and arithmetic solving strategies are appropriate. The following heuristic strategies are the most suitable: Guess – check – revise, Systematic experimentation, Solution drawing (see also Přibyl, Eisenmann, 2014).

There are commonalities as well as differences in Problem 9A and 9C statement structures. In both problems, the whole is known, there are three parts and one relationship between parts is multiplicative, the other additive. Part 2 is expressed in terms of Part 1 and Part 3 in terms of Part 2. There are two main differences between Problem 9A and 9C statements:

- In Problem 9A, both relationships are expressed by "more than". In Problem 9C, the additive relationship is expressed by "fewer than". Several studies suggest that the word "more" is comprehended easier than the word "less" (Riley and Greeno, 1988).
- Problem 9C can be solved either on the level of (1) the number of tours or (2) the earned CZK. In both cases, additional calculations must be done. (1) – the total number of tours must be calculated from the total amount of money and the price of one tour. (2) – in calculations, the money awarded for one tour must be used.

Design of test booklets

As stated above, each problem is in four variants that differ from each other in certain parameters. Also the sequence of problems in the test booklet (the order of problems and the connection to previous problems) can affect pupils' performance (both the success rate and the selected solving procedure).

The requirement was to make the variants of the test booklets relatively equally difficult and to respect general recommendations for creation of tests, i.e. that problems at the beginning of the test should be relatively less difficult. Also the time demandingness of all test booklets should be the approximately the same. That is why the individual problems had been piloted before the main testing. In the piloting, their time demandingness, their difficulty, presence of pupils' mistakes were studied. Based on these findings, we stated the total number of problems to be included in the test to make sure that ordinary pupils (not special needs pupils and pupils whose mother tongue was not Czech) could solve the test in 40 minutes without feeling any time pressure. The test booklets for the 9th grade, which included the problems analysed in this paper, consisted of the total of 6 problems.

When planning administration of the test booklets, we had to take into account that the variants of one problem are very similar at the first sight and always have the same result. That is why each pair of pupils sitting in one desk needed such a variant of the test booklet that looked very different from the other.

There were relatively many demands on the design of the test booklets. Based on these demands the following design of test booklets was created trying to respect all the demands (Table 3). One of them was the pupils' seating plan during testing.

Each variant of a problem was in a different test booklet. This implies that four different text booklets were created. Other four variants were created by changing the order of problems, see Tables 4 and 5.

One of the variants of each problem is the basic one. Changes in parameters usually make this basic problem cognitively more demanding. Sometimes a change of parameters makes the problem wording longer. That is why none of the test booklets could consist only of basic variants (i.e. A1, B1 etc.).

Position of the problem in the test \ number of test booklet	**T1**	**T2**	**T3**	**T4**
Problem 1	A1	A2	A3	A4
Problem 2	B4	B1	B2	B3
Problem 3	C3	C4	C1	C2
Problem 4	D2	D3	D4	D1

Table 3: Design of text booklets

With respect to the fact that the impact of a particular parameter may be affected by the position of the problem in the test booklet, for each of the four test variants a complementary variant was created. The couple of complementary variants was labelled I and II, where II means inverse order of problems. This means that the total of eight variants of test booklets was created (Table 4). Problems A and D were problems that came out as easier in piloting.

	T1_I	**T1_II**	**T2_I**	**T2_II**	**T3_I**	**T3_II**	**T4_I**	**T4_II**
Problem 1	A1	D2	A2	D3	A3	D4	A4	D1
Problem 2	B4	C3	B1	C4	B2	C1	B3	C2
Problem 3	C3	B4	C4	B1	C1	B2	C2	B3
Problem 4	D2	A1	D3	A2	D4	A3	D1	A4

Table 4: Variants of test booklets

Namely in test booklets for the 9th grade, the sequence of problems is given in Table 5. Bold is used for problems we focus on in this study.

	T1_I	T1_II	T2_I	T2_II	T3_I	T3_II	T4_I	T4_II
Problem 1	8B1	9D2	8B1	9D3	8B2	9D4	8B2	9D1
Problem 2	**9A1**	9B3	**9A2**	9B4	**9A3**	9B1	**9A4**	9B2
Problem 3	**9C4**	8D1	**9C1**	8D2	**9C2**	8D1	**9C3**	8D2
Problem 4	8D1	**9C4**	8D2	**9C1**	8D1	**9C2**	8D2	**9C3**
Problem 5	9B3	**9A1**	9B4	**9A2**	9B1	**9A3**	9B2	**9A4**
Problem 6	9D2	8B1	9D3	8B1	9D4	8B2	9D1	8B2

Table 5: Variants of test booklets for 9th grade

The basic scheme of the seating plan following the rule "maximum difference between test booklets of pupils sitting next to each other" is in Table 6:

T1_I	T3_II
T2_I	T4_II
T3_I	T1_II
T4_I	T2_II

Table 6: Seating plan when solving the tests

Problems 8B and 8D were used as anchor problems in IRT (Item Response Theory) analyses (Lord, 1980; Van der Linden and Hambleton, 1997). Other means of anchoring were the results from initial tests and assignment of test booklets to pupils in two subsequent rounds of testing in the GAČR project, which is shown in the following Table 7. Our problems were included in testing 1. Testing 2 consisted of different problems than testing 1.

	Pupils															
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
Initial test from Czech	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Initial test from mathematics	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Testing 1	**T1**	**T1**	**T1**	**T1**	**T2**	**T2**	**T2**	**T2**	**T3**	**T3**	**T3**	**T3**	**T4**	**T4**	**T4**	**T4**
Testing 2	T1	T2	T3	T4	T1	T2	T3	T4	T1	T2	T3	T4	T1	T2	T3	T4

Table 7: Assigning test booklets to pupils

Methods of analysis

To answer the research questions, quartets of Problem 9A and 9C were posed. The study is of a mixed methodology design, consisting of quantitative and qualitative parts.

Quantitative analysis

IRT was used both for the division of the pupils into equally able groups, each of which was solving a different variant of the word problem, and for the quantitative interpretation of data. The pupils' written solutions were analysed. The scoring rubric was as follows: 0 points (no or wrong solution), 1 point (partially correct solution), 2 points (correct problem model with a numerical mistake), 3 points (correct solution).[1]

To analyse the parameters of problems, a two-parameter logistic

1 The aim was to assess pupils' understanding of the problem, not their ability to carry out numerical operations and thus we considered as correct all solutions in which an appropriate problem model was created.

model was used (Lord, 1980): $P_{ij} = \frac{1}{1+e^{-a_i(\theta_j - b_i)}}$, where P_{ij} is the probability that a pupil j with the ability of Θ_j will solve the problem i correctly, a_i characterises problem discrimination and b_i its difficulty. To find the latent ability of pupils, a scale was put on a z-score with Bayes estimation EAP, in an iterative way using results in the Initial Test (for which a model graded for the total result in the test was used) and Test 1 (9A) and Test 2 (9C). This allowed compensation for any inconsistencies in terms of ability grouping which may have arisen at the beginning.

The situation is visually depicted in Figure 1. Θ_j is on the *x*-axis and the value of 0 means an average result in the initial test. On the *y*-axis we can read the probability with which the pupil of a certain ability will solve the task. For example, an average pupil will solve 9A2 with the probability approximately 0.6 (Figure 1 left), 9C2 with the probability approximately 0.14 (Figure 1 right).

Qualitative analysis

Finally, a qualitative analysis of the data was carried out. The pupils' written solutions were analysed carefully for mistakes and solving strategies with the main focus on the relationships between the variants of the problem and order of information recorded in pupils' legends. In this case, we created a spreadsheet in which the order of information in the legends was recorded for each variant of the two problems. The occurrence of frequent mistakes was linked with the use of the order of information in legends.

Results

In this section, we present the results gained from our experiments. The results are divided into sections focusing on quantitative and qualitative analyses.

Quantitative analysis using IRT

Table 8 presents the relative frequencies of point distributions and the average success rate for the four variants of Problem 9A. It seems that most difficult are the variants 9A4 and 9C4, while in the variants 9A1, 9A2, 9A3, and 9C1, 9C2, 9C3 average success rates are similar. The order of numerical data played its role in case of a less familiar context, resp. formulation without "if". This corresponds to our expectations. The results confirm that Problem 9C was more difficult than 9A in three variants while in variants 9A4, 9C4 the average success is comparable and low, which was surprising. This can be explained by the fact that in both cases the pupils came across some "intricacy" that made the use of school procedures more difficult. These results are further specified by IRT evaluation and verified by the statistical methods below.

Table 8 and Figure 1 present the parameters of the problems according to IRT. They add more detail to the information that we can read from the average success rate. In Figure 1, graphs on the left in suggest that 9A2 has the best discrimination among 9A variants. 9A1 and 9A2 discriminations are comparable. Below-average pupils solved the variant with an unfamiliar context better, while the average and above-average pupils were not much affected by the context. 9A3 discriminates the worst pupils. Graphs on the right suggest that 9C is much more difficult than 9A – an important number of above-average pupils failed to solve its variants. 9C4 has the best discrimination. 9C2 and 9C3 discriminations are comparable.

	N	*Points* 0	1	2	3	*Average success rate*		*N*	*Points* 0	1	2	3	*Average success rate*
	Familiar context							Unfamiliar context					
9A1	46	37%	9%	2%	52%	57%	**9A3**	45	29%	18%	7%	46%	57%
9A2	46	28%	15%	9%	48%	59%	**9A4**	45	36%	20%	7%	37%	49%
	Formulated with "if-clause"							Formulated without "if-clause"					
9C1	45	42%	33%	7%	18%	33%	**9C3**	46	46%	26%	11%	17%	33%
9C2	46	30%	43%	7%	20%	38%	**9C4**	45	51%	38%	4%	7%	22%

Table 8: Results for 9A, 9C according to success rate

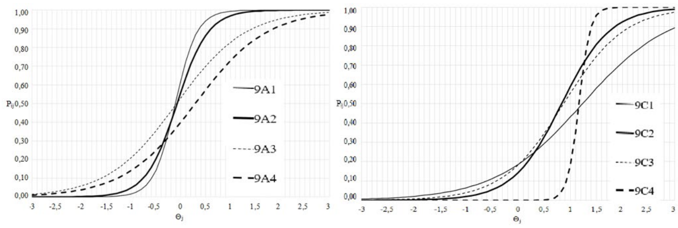

Figure 1: Two-parameter IRT model for the variants of 9A (left) and 9C (right)

Qualitative analysis

Solving strategies

The experiment was conducted with 182 pupils. Only 12 of them selected successfully as their solving strategy the use of an equation (7 used Guess – check – revise: 9A1 – 1 pupil, 9C1 – 1 pupil, 9C2 – 1 pupil, 9C3 – 3 pupils, 9C4 – 1 pupil). 6 pupils solved the problem arithmetically as a problem on division of the whole into three equal parts, which can be accounted for by their failure to understand the conditions in the text statement correctly (9A3 – 1 pupil, 9A4 – 1 pupil, 9C1 – 1 pupil, 9C2 – 2 pupils, 9C3 – 1 pupil).

Legends

In Problems 9A and 9C, the proper order of numerical data is (P1, P2, P3, W), where W describes the whole in all variants. In 9A1, 9A2 P1 refers to gymnastics, P2 basketball and P3 swimming whereas in 9A3, 9A4 P1 refers to the Liberal Club, P2 to the Conservative Club and P3 to the Green Club; in 9C P1 stands for Eva's, P2 for Karel's and P3 for David's number of tours. Table 9 and 10 present the distribution of direct and indirect legends in comparison to the proper order of numerical data in the statement. Recorded are only those cases where the pupil attempted to solve the problem and created a legend (independently of whether they solved the problem successfully). In some cases, the pupil used the whole, but they did not record it in the legend. The order of parts in legends that did not appear in the pupils' solution are not included in tables 9 and 10.

Variant	Order of numerical data in the statement	Legend: Position of W	Legend: Order of parts
9A1	**P2, P1, P3, W**	***W as last: 30***	***P2, P1, P3: 11***
		W as first: 4	P1, P2, P3: 20
			P2, P3, P1: 3
			P3, P2, P1: 0
9A3	**P2, P1, P3, W**	***W as last: 25***	***P2, P1, P3: 29***
		W as first: 2	P1, P2, P3: 1
		Not recorded: 7	P2, P3, P1: 4
			P3, P2, P1: 0
9A2	**W, P2, P1, P3**	W as last: 18	***P2, P1, P3: 11***
		W as first: 15	P1, P2, P3: 22
		Not recorded: 3	P2, P3, P1: 2
			P3, P2, P1: 1
9A4	**W, P2, P1, P3**	W as last: 17	***P2, P1, P3: 26***
		W as first: 9	P1, P2, P3: 0
		Not recorded: 4	P2, P3, P1: 3
			P3, P2, P1: 1

Table 9: Distribution of direct and indirect legends for 9A

Table 9 shows that in case of proper order of numerical data in the statement most pupils use direct legend independently of familiarity of context. In case of mixed order of numerical data in the statement the situation is different for familiar and unfamiliar contexts: In case of an unfamiliar context, the proportion of pupils who changed the order of information in legend, albeit by moving W to the end or by swapping the order of recorded parts is much higher than in case of familiar context. As stated earlier in the paper, Problem 9C is more difficult than Problem 9A as it does not contain only the relationship of the whole to parts, but also the relationship between the money earned and number of tours. Thus there are two possible approaches to the solution. One is to find out how many tours were taken, find the number of tours of each of the students using the information from the problem statement and finally calculate the money earned by each student. The other is to work directly with earned money, then it is essential to supplement the relationship between parts on the award for one tour; and the results are in crowns. If pupils do not realize what approach they use, they may come to a wrong result.

Variant	Order of numerical data in the statement	Legend: Position of W	Legend: Order of parts
9C1	**W, P2, P1, P3**	***W as first: 7***	***P2, P1, P3: 29***
		W as last: 21	P1, P2, P3: 3
		Not recorded: 4	P2, P3, P1: 0
Number of pupils who interchanged level incorrectly: 11			
9C3	**W, P2, P1, P3**	***W as first: 10***	***P2, P1, P3: 22***
		W as last: 13	P1, P2, P3: 2
		Not recorded: 5	P2, P3, P1: 4
Number of pupils who interchanged level incorrectly: 8			
9C2	**P2, P1, P3, W**	W as first: 2	***P2, P1, P3: 28***
		W as last: 22	P1, P2, P3: 2
		Not recorded: 8	P2, P3, P1: 2
Number of pupils who interchanged level incorrectly: 21			
9C4	**P2, P1, P3, W**	W as first: 4	***P2, P1, P3: 29***
		W as last: 25	P1, P2, P3: 2
		Not recorded: 4	P2, P3, P1: 2
Number of pupils who interchanged level incorrectly: 25			

Table 10: Distribution of direct and indirect legends for 9C

Table 10 is analogical to Table 9 for Problem 9C. In addition the table shows how many pupils made mistakes when switching between the levels of the number of tours and money earned. Table 10 shows that most pupils used direct legend in variants 9C2 and 9C4; this can be explained by the order in which data are used in calculations. In the variants 9C1 and 9C3 the order of parts mostly corresponded to the problem statement but the information on the whole was often in a different place that in the statement. This could be caused by the two-level problem statement which requires from pupils to pay a lot of attention to all data from the beginning in order to understand and by their attempt to find out the total number of tours. The table also shows that in the variants with W at the end of the statement independently of the presence of "if-clause" pupils make more mistakes in switching between the number of tours and money earned. This could be explained by the fact that they can process the relationships between parts without paying attention to the money issue.

Analysis of mistakes that occurred in the solutions of problem 9A

The structure of problem 9A can be described by the equation $x + 3x + (3x + 114) = 380$. The problem can be solved in several ways, not only using an equation, but the 9th graders can be expected to be so familiar with linear equations that they will prefer this way of solving the problem.

Table 11 presents the mistakes pupils made when solving the problem. It does not contain erroneous solutions where the used solving procedure could not be identified (3 cases in 9A2, 1 case in 9A3 and 9A4). Apart from *Division of the whole into three parts* (procedure I), the given procedures in the variants were used only by one or two pupils. That is why we do not present the numbers in brackets in case of solving procedures II to XII.

I	*Division of the whole into three equal parts*	9A1 (3), 9A2 (4), 9A3 (3), 9A4 (3)
II	$x + 3x - 114 = 380$	9A1
III	$x + 3x + 3(x - 114) = 380$	9A1
IV	$x + x + 3x + (3x + 114) = 380$	9A2
V	$3x + x = 114 \text{ x } 3$	9A4
VI	$x + 3x + 114 = 380$	9A1, 9A4
VII	$3x + x + (x + 114) = 380$	9A1, 9A3, 9A4
VIII	$(380 - 114): 2$	9A4
IX	$3x + 114 = 380$	9A3, 9A4
X	$3x - 114 = 380$	9A3, 9A4
XI	$x + 3x + (3x + 114) = x$	9A2
XII	$3x - 114 + x + y = 380$	9A4

Table 11: Types of mistakes in variants of problem 9A

Table 11 shows that there were 12 different erroneous procedures in pupils' solutions. Only procedure I can be come across in all variants. Procedure VII was used in three variants, 3 procedures (VI, IX, X) were used in two variants.

The very list of erroneous procedures and variants in which they occurred shows clearly that pupils found variant 9A4 where 8 different wrong procedures were used the most difficult. In variants 9A3, 9A4, i.e. in variants with an unfamiliar context, the same erroneous procedure was used in two cases (IX, X). Procedures VI, VII can be come across once in variants with familiar context and once in a variant with unfamiliar context.

Figure 2 shows an example of solution using procedure I (division of the whole into three equal parts).

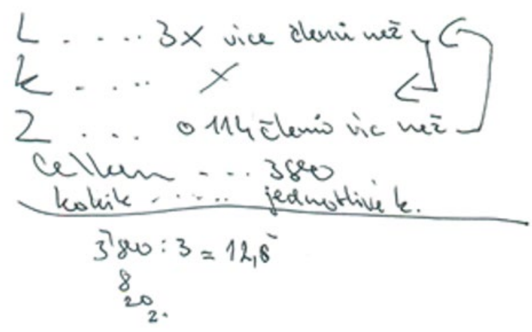

Figure 2: Variant 9A3, procedure I

Division of the whole into three equal parts is a mistake that is often to be come across in this type of problems. (Novotná, 2010) The strategy is used most often by those solvers who have problems with comprehending the text. Since they cannot grasp the relations between unequal parts, they look for salvation in division into equal parts.

The use of other erroneous procedures implies that the pupils did not give up on dividing the whole into unequal parts but failed to mathematize the conditions from the problem wording correctly. The number of solving procedures used signals that mathematization of conditions from the problem wording was far from easy. Figures 3 and 4 show examples of pupils' solutions of variant 9A4 and illustrate how the pupil was trying to handle all relevant conditions unsuccessfully. Figure 5 shows a solution where the pupil managed to find the correct solution after several unsuccessful attempts.

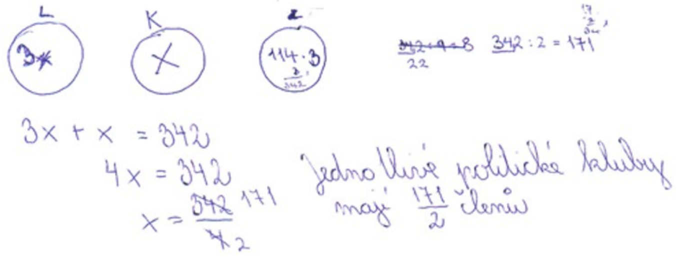

Figure 3: Variant 9A4, procedure V (in Czech 114. 3 stands for 114 x 3)

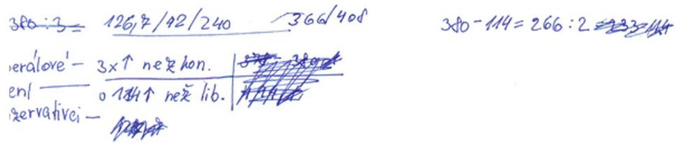

Figure 4: Variant 9A4, several unsuccessful attempts without reaching the correct procedure

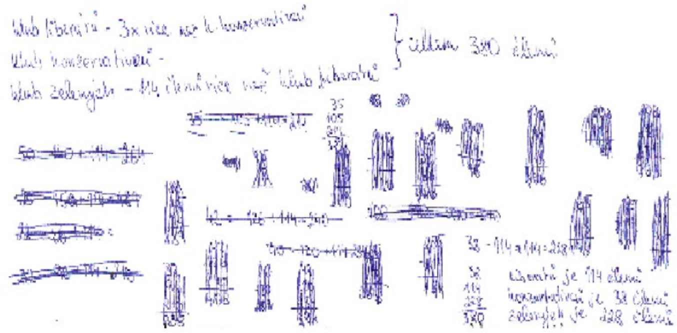

Figure 5: Variant 9A3, the correct procedure reached after several unsuccessful attempts

The last row in Table 11 shows the use of two unknowns in the legend and solution (see Figure 6). This is probably the result of the pupil's failure to mathematize all the relations between the sought values, in consequence of which he tried to map the situation using two different letters for two different unknowns. This is not a mistake. The problem is the pupil did not manage to express the relation between the two sought values correctly.

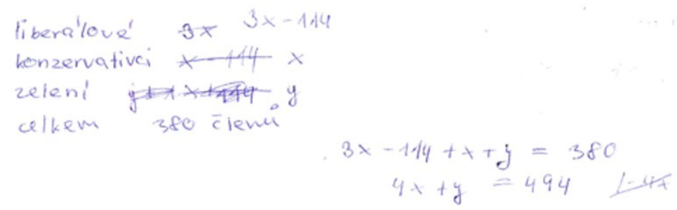

Figure 6: Variant 9A4, procedure XII

Table 12 shows the number of pupils who did not even start solving problem 9A.

Variant 9A1	Variant 9A2	Variant 9A3	Variant 9A4
10	4	10	7

Table 12: Numbers of pupils who did not solve problem 9A

Table 12 shows that the order of information had major impact on this indicator. As we mentioned already in the description of the foursome of variants of problem 9A, the order of information in the problem wording was the same in the couples of variants 9A1, 9A3 and 9A2, 9A4. At the same time the number of pupils who did not start solving the problem in variants 9A1, 9A3 is significantly higher than in variants 9A2, 9A4. It seems the consequence of different order of information presented in the wording of problems 9A1, 9A3 was that these problems were much more difficult for pupils than variants 9A2, 9A4. A higher number of pupils who did not start solving in variant 9A4 than 9A2 implies that the unfamiliar context could have been relevant. However, in the research sample the same does not hold for variants 9A1, 9A3, where the number of pupils who did not start solving is balanced.

Discussion

This study of the impact of the order of presenting data in the wording of a mathematics problem brings new knowledge into the area of word problem solving. It comes out of an analysis of pupils' written solutions exclusively. Thus, the causes of the mistakes remain unclear in many cases, as written solutions do not provide enough information. If we want to get information on the impact of the order of data that is more precise, we should supplement the analysis of pupils' written solutions with interviews with the pupils. So far, several interviews with pupils in which the tested problems were discussed have been conducted with pupils not involved in the testing. These interviews reinforce the results gained from the testing. The gained results comply with the results in (Vicente, Orratia, Verschaffel, 2008), where a significant impact of the order of information in wording of additive problems with two transformations was detected.

Downton and Sullivan (2017: 303) indicate that 'posing of appropriately complex tasks may actually prompt the use of more sophisticated strategies'. This was not the case of Problems 9A and 9C. The success rate in Problem 9C was smaller than in Problem 9A (see Table 8). This can be explained by a higher level of complexity of Problem C. The used strategies were comparable.

An important issue is the dependence of the order of information in the wording of a problem on the solvers' age. In the project GAČR 16-06134S *Context problems as a key to the application and understanding of mathematical concepts*, the impact of the order of information is analysed with pupils from several age categories. So far, unpublished analyses show that the impact of the order of information decreases with growing age of the solvers. For example, analyses of solutions of 4th graders show that the differences when changing the order of data are statistically significant. This does not hold for 9th graders, which can be accounted for both by pupils' greater experience with problem solving, by their higher proficiency in reading but also by their use of more advanced mathematical solving strategies, namely equations. Setting up an equation or a set of equations describing the structure of the problem wording is characterized by parallel grasping of information from the wording (Novotná, 2010) when the solver first registers all relevant information and then creates the appropriate mathematical model. Then the order of information becomes less significant for successful solution of the problem.

As shown in the analysis of legends created by the tested pupils, not all pupils make a legend while solving. In fact, creation of a legend is coding of data from the problem wording in a new form that is more comprehensible or clearer to the pupil. Gelfman and Kholodnaya (1999) consider ways of coding data to be a subjective means that help us project the world around us into our own experience. This also holds for processing of data from the word problem wording. If the pupil is convinced that they grasp the structure of the problem, they do not need to visualize it and do not make a legend. However, if the structure of the problem wording is harder to grasp for any reason, creation of a legend may help the pupil get insight into the structure of the problem. A well-created legend can make solving of a problem easier for the solver, an inconveniently selected legend can make the solution much harder (Novotná, 2010).

Conclusions

The variables analysed here for 14-16-year old pupils are studied for pupils 8 to 16 in the GAČR project. The research in the GAČR project has important characteristics. The variations of the problems are carefully formulated so that only one variable was changed, the other being controlled (as requested by Daroczy et al. (2015)). Great care is paid to dividing pupils into equally able groups using IRT. The quantitative analysis is accompanied by a qualitative analysis to better understand the influence of investigated variables.

Analogical analyses of pupils' solutions in which various mathematical, psychological and language variables are changed allow us to get deeper insight into pupils' difficulties when solving word problems. The presented study has its limitations: a relatively small number of participants, one type of word problems, analyses based on written solutions only. However, it brought very interesting results important for mathematics education as well as test creation. The study brings more light to the role of word problems in mathematics education. As Mason (2018: 334) states: [Pupils should be taught to] 'recognise and express relationships in the situation using whatever support devices and modes of presentation are recommended for this purpose'.

Understanding the impact of the order of data on the difficulty of problems is important not only for teachers but also for authors of tests, curricular documents and textbooks. If problems with *mixed* order of data are used in teaching, due attention must be paid to efficient strategies allowing pupils to understand relations among them. Authors of tests as well as others should be aware of the fact that a change in the order of data in the problem wording might be seemingly small but might affect the difficulty of the problem and its discriminative property with an impact on the test reliability significantly.

Acknowledgement

This research was financially supported by the grant GACR 16-06134S Context problems as a key to the application and understanding of mathematical concepts.

References

Daroczy, G., Wolska, M., Meurers, W. D. and Nuerk, H. C. (2015) 'Word problems: a review of linguistic and numerical factors contributing to their difficulty', *Frontiers in Psychology*, vol. 6, no. 348, pp. 22–34. http://dx.doi.org/10.3389/fpsyg.2015.00348

Downton. A. and Sullivan, P. (2017) 'Posing complex problems requiring multiplicative thinking prompts students to use sophisticated strategies and build mathematical connections', *Educational Studies in Mathematics*, vol. 95, no. 3, pp. 303-328. http://doi.org/10.1007/s10649-017-9751-x

Gelfman, E. and Kholodnaya, M. (1999) 'The role of ways of information coding in student's intellectual development', *Proceedings of the First Conference on the European Society for Research in Mathematics Education*, vol. 2, pp. 38-48, [Online], Available: http://www.mathematik.uni-dortmund.de/~erme/doc/cerme1/cerme1_proceedings_part2.pdf. [15 Nov 2018].

Hembree, R. (1992) 'Experiments and relational studies in problem solving: A meta-analysis', *Journal for Research in Mathematics Education*, vol. 23, no. 3, pp. 242-273. http://dx.doi.org/10.2307/749120

Lord, F. M. (1980) *Applications of item response theory to practical testing problems*, Hillsdale, N.J.: Lawrence Erlbaum Associates.

Mason, J. (2018) 'Structuring structural awareness: A commentary', *Building the foundation: Whole Numbers in the Primary Grades,* Cham: Springer International Publishing, pp. 323-340.

Nesher, P., Hershkovitz, S. and Novotná, J. (2003) 'Situation model, text base and what else? Factors affecting problem solving', *Educational Studies in Mathematics*, vol. 52, no. 2, pp. 151-176. https://doi.org/10.1023/A:1024028430965

Novotná, J. (2010) *Study of solving word problems in teaching of mathematics. From atomic analysis to the analysis of situations*, Saarbrücken: LAP LAMBERT Academic Publishing.

Novotná, J. (2018) 'Impact of order of data in word problems on division of a whole into unequal parts', *ERIE 2018*, Prague: CULS, pp. 257-264.

Novotná, J., Eisenmann, P., Přibyl, J., Ondrušová, J. and Břehovský, J. (2013) 'Heuristic strategies in problem solving in school mathematics', *ERIE 2013*, Prague: CULS, pp. 461-468.

Novotná, J. and Vondrová, N. (2017) 'Pupils' strategies for missing value proportional problems', *Proceedings of ERIE 2017*, Prague: CULS, pp. 279-286.

Přibyl, J. and Eisenmann, P. (2014) 'Properties of problem solving strategies', *Proceedings of ERIE 2014*, Prague: CULS, pp. 623-630.

Riley, M. S. and J. G. Greeno (1988) 'Developmental Analysis of Understanding Language about Quantities and of Solving Problems', *Cognition and Instruction*, vol. 5, no. 1, pp. 49-101. http://dx.doi.org/10.1207/s1532690xci0501_2

Searle, B.W., Lorton, P. Jr. and Suppes, P. (1974) 'Structural variables affecting CAI performance on arithmetic word problems of disadvantaged and deaf students', *Educational Studies in Mathematics*, vol. 5, pp. 371-384, http://dx.doi.org/10.1007/BF01424555

Soneira, C., González-Calero, J.A. and Arnau D. (2018) 'An assessment of the sources of the reversal error through classic and new variables', *Educational Studies in Mathematics*, vol. 99, no. 1, pp. 43-56. https://doi.org/10.1007/s10649-018-9828-1

Van der Linden, W.J. and Hambleton, R.K. (1997) *Handbook of Modern Item Response Theory,* New York: Springer, http://dx.doi.org/10.1007/978-1-4757-2691-6

Vicente, S., Orrantia, J. and Verschaffel, L. (2008) 'Influence of situational and mathematical information on situationally difficult word problems', *Studia Psychologica*, vol. 50, no. 4, pp. 337-356.

Vondrová, N., Novotná, J. and Havlíčková, R. (2019) 'The influence of solving and situational information on pupils' achievement in additive word problems with several states and transformations', *ZDM Mathematics Education,* pp. 1-15, https://doi.org/10.1007/s11858-018-0991-8

Wijaya, A., van den Heuvel-Panhuizen, M. And Doorman, M. (2015) 'Opportunity-to-learn context-based tasks provided by mathematics textbooks', *Educational Studies in Mathematics*, vol. 89, no. 1, pp. 41-65. https://doi.org/10.1007/s10649-015-9595-1

Zohar, A. and Gershikov, A. (2008) 'Gender and Performance in Mathematical Tasks: Does the Context Make a Difference?', *International Journal of Science and Mathematics Education*, vol. 6, no. 4, pp. 677-693, http://dx.doi.org/10.1007/s10763-007-9086-7

EFFICIENCY AND RESPONSIBILITY IN EDUCATION ERIE 2019

Conference Efficiency and Responsibility in Education 2019 is an international conference that is organized by the Faculty of Economics and Management, Czech University of Life Sciences Prague on June 6 - 7, 2019.

Conference topics

- Theory and Methodology of Pedagogy and Education
- Information and Knowledge in Lifelong Education and Training
- ICT in Education
- Applications, Practice and Experience
- Responsibility in Education, Ethical Issues
- Generation Z: Labour market perspectives

All papers are submitted to a double-blind peer review process before publication. Based on reviewers' recommendation, the papers will be published in the conference proceedings.

The previous conference proceedings are listed on the Thomson Reuters ISI Index to Social Sciences & Humanities Proceedings (ISSHP) and the Thomson Reuters ISI Index to Social Sciences & Humanities Proceedings (ISSHP/ISI Proceedings).

Deadline for full paper submission is February 24, 2019.

Programme Committee

- Head: Jaroslav Havlíček, Czech University of Life Sciences Prague, Czech Republic
- Members:
    - Peter M. Bednar, University of Portsmouth, United Kingdom
    - Peter Fandel, Slovak University in Nitra, Slovakia
    - Jana Hančlová, Czech Society for Operations Research, Czech Republic
    - Milan Houška, Czech University of Life Sciences Prague, Czech Republic
    - Irem Comoglu, Dokuz Eylul University, Turkey
    - Piet Kommers, University of Twente, Netherlands
    - Eva Milková, University of Hradec Králové, Czech Republic
    - William O'Brien, Worcester State University, United States
    - Karel Sladký, Academy of Sciences of the Czech Republic, Czech Republic
    - Milan Turčáni, Constantine the Philosopher University in Nitra, Slovakia
    - Jarmila Novotna, Charles University, Czech Republic

Editors

- Jiří Fejfar, Czech University of Life Sciences Prague
- Martin Flégl, La Salle University, Mexico City
- Martina Fejfarová, Czech University of Life Sciences Prague
- Hana Urbancová, Czech University of Life Sciences Prague
- Jakub Husák, Czech University of Life Sciences Prague
- Milan Houška, Czech University of Life Sciences Prague
- Igor Krejčí, Czech University of Life Sciences Prague

Conference co-editors

- Irena Benešová, Department of Economics, CULS Prague
- Jiří Brožek, Department of Information Engineering, CULS Prague
- Helena Brožová, Department of Systems Engineering, CULS Prague
- Ludmila Dömeová, Department of Systems Engineering, CULS Prague
- Robert Hlavatý, Department of Systems Engineering, CULS Prague
- Tomáš Hlavsa, Department of Statistics, CULS Prague
- Martina Houšková Beránková, Department of Systems Engineering, CULS Prague
- Hana Chýlová, Department of Psychology, CULS Prague
- Petr Kučera, Department of Systems Engineering, CULS Prague
- Martina Kuncová, Department of Economic Studies, College of Polytechnics Jihlava
- Ludmila Natovová, Department of Psychology, CULS Prague
- Josef Pavlíček, Department of Information Engineering, CULS Prague
- Jan Rydval, Department of Systems Engineering, CULS Prague
- Petra Šánova, Department of Trade and Accounting, CULS Prague

- Pavel Šimek, Department of Information Technologies, CULS Prague
- Eva Vaněčková, University of South Bohemia, Czech Republic
- Kristýna Vltavská, Department of Economic Statistics, University of Economics, Prague

Conference Fee

- Full conference fee: 150 EUR or 3800 CZK
- Reduced conference fee (for regular PhD students): 90 EUR or 2200 CZK
- Surcharge for additional page: EUR 20 or CZK 500 per additional page (more than 7 pages)
- Participant´s accompany, co-authors (without conference materials): 40% discount of the conference fee
- Participant´s accompany, co-authors (with conference materials): full price of the conference fee

Conference deadlines

- November 20, 2018 - 1st call for papers
- January 7, 2019 – 2nd call for papers
- **February 24, 2019 – Deadline for full paper submission**
- April 1, 2019 – Results of review process
- April 21, 2019 – Deadline for final version submission
- May 12, 2019 – Notification of paper acceptance
- May 26, 2019 – Deadline for conference fee payment

For more information do not hesitate to contact the conference editors at erie@pef.czu.cz or visit the conference website https://erie.pef.czu.cz/en/.