

EFFICIENCY OF KNOWLEDGE TRANSFER THROUGH KNOWLEDGE TEXTS: STATISTICAL ANALYSIS

Tereza Rauchová, Milan Houška

Czech University of Life Sciences Prague

rauchova@pef.czu.cz

Abstract

Texts are an important way to share and transfer knowledge. In this paper we analyse the impact of a specific form of texts, so called “knowledge texts”, on the efficiency of knowledge transfer. The objective is to verify or reject several hypotheses on the relationships among the style of educational texts (standard or knowledge styles), learning outcomes (performance of the students after learning) and subjective evaluation of conformity of working with individual styles of the texts. For this purpose, we carry out experiment with a homogeneous group of the students ($n = 41$) divided into an experimental group and a control group. We use statistical methods to process the results of the experiments; ability of the students to solve specific tasks and their opinions on readability and understandability of the texts subject to the time spent for learning. Even if we determine statistically significant relationships between the style of texts and accuracy of the problem solving in the experimental group only, the results allow us to improve the experiment and apply the methodology developed in a less structured branch than the Operational Research (Graph Theory) is. The methodology is another benefit of the paper, because it can be applied independently on a particular domain.

Key Words

Knowledge unit, Standard text, Knowledge text, Statistical analysis, Graph theory

ARTICLE INFO

Article type

Full research paper

doi: 10.7160/eriesj.2013.060105

Article history

Received: March 8, 2012

Received in revised form: March 29, 2012

Accepted: March 30, 2013

Available on-line: March 31, 2013

Introduction

Studies on the role and efficiency of educational texts as a form of knowledge transfer are very topical, particularly in pedagogical sciences. Pedagogical researchers are very often focused on measuring the efficiency of education (Tudor, 2012). Teaching methods are an important issue in measuring their efficiency (Maňák, Janík 2009; Starý, Chvál, 2009). Vališová and Kasíková (2011) describe several teaching methods. Working with text is one of them. Starý and Chvál (2009) describe models of quality and efficiency on pedagogical level. However, these models have different characteristics from the systems approach models; e.g. Data Envelopment Analysis (Šubrt, 2011). Dömeová et al. (2008) use the systems approach for knowledge modelling. Also this approach is used by Glava and Glava (2011); they combine didactic modelling with an analytic point of view.

Teaching texts and textbooks are used in many cases during lessons but also during home studying and preparing. Texts and textbooks have many significant functions and they have to be analyzed in detail. For this reason, experiments are often used in order to clarify the importance of some properties and parameters of textbooks (Mikk, 2007). Tannenbergová (2009), Dobrylovský (2009) or Hodis (2003) focused on the analysis of the pedagogical texts. The observed aspects are, for example, the measuring of difficulty, analysis of terminology, didactic content of text, or information density, etc. These aspects are measured and expressed by quantifiers.

Prasad and Ojha (2012) present an experiment on comparing three ways of transferring knowledge (text, table and graphs) and then evaluating their efficiency. They use the speed of understanding and accuracy of responses as criteria. Based on their experimental data they discovered that there is no ideal

form of knowledge transferring because of antagonistic criteria. The fastest way (the graphs) is the least accurate one and vice versa, the most accurate way (the text) is the slowest one.

Kools et al. (2006) deal with a similar problem; how to evaluate the effect of graphic organizers on the comprehension of a specific educational text and compared subjective with objective comprehension measures. They found significant positive impacts of graphic organizers at four levels of objective comprehension as indicated by open comprehension questions. Obviously, comprehension is also influenced by the graphical way of presentation of knowledge in texts.

Lee and Segev (2012) stress the impact of a specific form of declarative representation of knowledge; knowledge maps (K-maps) in learning and e-learning. In contrast with the traditional approaches to the construction of such maps by human experts, they propose a text mining technique to create it automatically. For this purpose they use the TF/IDF algorithm to extract keywords and then they develop the K-maps in the current domain based on the ranking pairs of keywords according to the number of their occurrences in a sentence and the number of words in a sentence. Auxiliary experiments show that the K-map learning identifies core ideas much more smoothly compared to the standard document learning. Finally, the K-maps are a promising and commonly used tool in more areas of working with knowledge and people, e.g. in human resources management or identification of talents (Kolman, 2008).

The objective of the paper is to analyze the impact of different styles of educational texts on the performance provided by the students. In particular, we focus on two ways of the text presentations:

- standard text, as usually appears in textbooks;
- knowledge text, redesigned by using the methods of Knowledge Engineering.

We concentrate on accepting or rejecting the following main hypotheses:

The students provide significantly better performance using the knowledge text rather than the standard one. They solve the problems more successfully. Also, the style of the knowledge text is more comfortable for them to work with.

We carried out an experiment with the students of the Czech University of Life Sciences Prague. Two groups of students worked with the standard and knowledge educational texts and solved problems in the area of Mathematical Methods in Economics. Objective results and subjective evaluation of the texts received from the students were processed by advanced statistical methods. That allowed us to decide on the validity of the above-mentioned hypotheses.

Materials and Methods

Knowledge texts

In this study, we understand the “knowledge text” as a specific form of educational text, which contains knowledge in an explicit form.

In particular, apart from the necessary data and information, there is also procedural knowledge (usually represented by a production rule or a knowledge unit). The knowledge is expressed in a natural language in the knowledge text.

Dömeová, Houška and Beránková (2008) suggested a definition of the “knowledge unit” (KU) as a special, well-structured

type of a piece of knowledge that contains one production rule related to the successful solving an elementary problem. Formally, a knowledge unit can be expressed as

$$KU = \{X, Y, Z, Q\}, \quad (1)$$

where X stands for a problem situation,

Y stands for an elementary problem being solved within the framework of the X problem situation,

Z stands for an objective of solving the elementary problem,

Q stands for a successful solution of the elementary problem (result).

The elementariness of knowledge is predetermined by the elementariness of the problem. The elementary problem is a problem or a part of a complex problem which is impractical to be further divided into more simple sub-problems. Criteria for assessing the degree of elementariness are defined by the knowledge user, because they depend on his or her ability to understand and apply the rules included in elementary knowledge. This is in conformity with Zack’s definition of knowledge units (Zack, 1999).

Knowledge unit may be expressed as a whole in a natural language. As mentioned above, there is no exclusivity; each part of the unit has several facultative ways of expression and almost all of their combinations are feasible. The basic form of knowledge unit expression derived by systems approach is defined as follows:

If we want to solve an elementary problem Y in the problem situation X to reach the objective Z , then we should apply the solution Q .

To create the knowledge text from the standard one, we use the following procedure (Dömeová, Houška, Beránková, 2008).

- Step 1: Select a standard text for transformation.
- Step 2: Analyze the text to identify knowledge units included.
- Step 3: Express the knowledge units in an analytical form.
- Step 4: Convert the analytical form of knowledge unit into a natural language.

Research sample

The experiment was accomplished with 41 students in total. All participants study the programme Public Administration and Regional Development at the Czech University of Life Sciences Prague. The participants were divided into two groups: A (experimental group) and B (control group). The groups were balanced subject to the following criteria: age, gender, prior qualification, prior formal education and mathematical skills measured by the study results in the study subject Mathematical Methods in Economics and Management reached in the past.

Design of the experiment

The experiment is observed by the administrator who is responsible for ensuring the same conditions for all participants, recording the outputs from the participants and doing auxiliary tasks during the experiment (measuring the time, distributing data sheets and questionnaires, etc.). In case of doubts he answers the questions from the participants. The experiment was realized within two weeks.

Week 1. The group A receive the knowledge text, the group B the standard educational text to study the methods to solve the problem. The following aspects are observed:

- Duration of studying the texts measured by the time necessary to understand it. Due to psychological reasons (to avoid the

disturbance of the participants), this aspect is monitored covertly. The start is common at the same time, the end is announced by the student individually to the administrator.

- Quality of understanding measured by the ability of the students to solve a specific problem using their new knowledge. Two-value measure "Pass"/"Fail" is used.

Week 2. The group A receive the standard educational text, the group B the knowledge text to study the methods to solve the problem. Again, the texts deal with the same topic, but the particular algorithm differs from the one in the Week 1. Except Duration of studying the texts and Quality of understanding measured identically as in the Week 1, also we observe subjective opinions of the students on the comfortableness of working with standard and texts. In further text, we use the variables denoted as follows.

- t_1 ... time of learning - week 1 (s)
- t_2 ... time of learning - week 2 (s)
- acc_1 ... accuracy of the problem solving (number of users solved the problem correctly) - week 1
- acc_2 ... accuracy of the problem solving (number of users solved the problem correctly) - week 2

The process of the experiment is depicted in Figure 1.

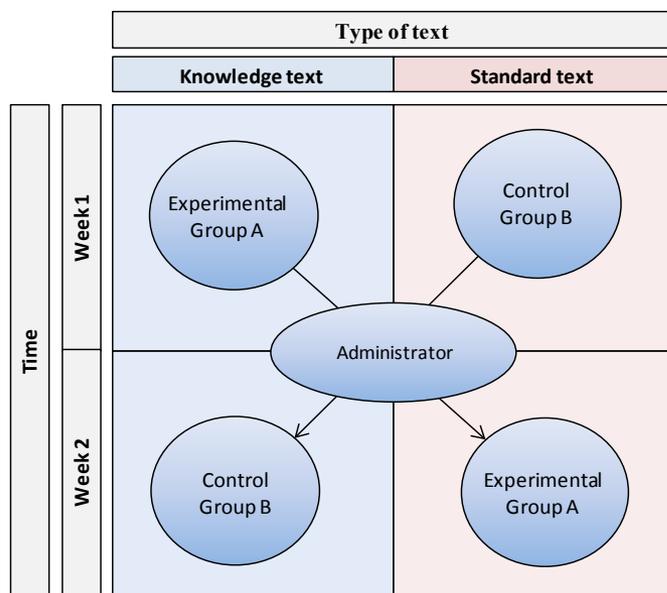


Figure 1: Design of the experiment

Statistical methods used

We use the following statistical methods to analysing the data received from the experiment. In details, all methods are described in statistical literature, e.g. Lindseuy (2009) or Peck and Devore (2012).

Basic descriptive statistics

The statistical characteristics are called numeric values that provide us the basic information about the statistical properties

of the population. For this work we use the characteristics that describe the measure the central tendency and dispersion of the data.

We use arithmetic mean, median, variance and standard deviation to describe the sample in our experiment. The mean value is observed because of the differences in duration of comprehension (measured in units of time) in normal and knowledge texts. Frequencies are used for description of execution, correctness and accuracy of the problems solved.

Parametric tests

Hypotheses are regarding the value of one or more parameters of the distribution. We assume a normal distribution of random tested variables. These tests are numerically difficult, but they have a good power of the test. Unknown parameter values between the two populations can be measured by two-sample parametric tests.

The experiment is based on two groups (A and B). The former works with knowledge text and the latter works with normal texts during the first part of the testing procedure. The role of the groups and type of texts are exchanged during the second part of the experiment (see Figure 1). Some chosen aspects are observed in each group. Therefore , two-sample F test, two-sample t-test, Behrens-Fisher test, Shapiro-Wilk W test and two-sample test about relative frequencies are all used in the experiment for finding differences between working with different type of texts.

Correlation analysis of qualitative variables

Contingency is a relationship of two or more quality characteristics, of which at least one is a sign of the plural. Characters can be organized into contingency tables. Each of

the characters is divided into k (rows) and m (columns) groups, where k is the number of permutations of the first character and m is the number of permutations of the second character. Chi-square test is used for testing the independence in the contingency table.

In our experiment we supplemented objective aspects also by subjective aspects. Subjective aspects are subjective expression and experience with solving the same problem. We analyze these aspects together and in association with execution, correctness and accuracy of the problems solved.

All calculations were carried out in the Statistica, version 10.

Results

We use the above-given statistical methods to find out, whether the users working with knowledge texts solve problems better than the others. Also we test the influence of the type of text on the performance of the users.

Week 1 - Data Analysis

Descriptive statistics - time of learning

We provide basic descriptive statistics for the variable t_1 for the experimental group A and control group B. The users from the group A work with the knowledge text, group B with the standard text. The statistics are summarized in Table 1.

	n	Mean	Median	Min	Max	Variance	Standard deviation
Group A	19	518.42	512.00	8.00	1014.00	79730.70	282.37
Group B	22	169.23	136.50	10.00	421.00	14459.90	120.25

Table 1: Descriptive statistics, variable: t_1 , week 1

Obviously, the mean value as well as the variance and the maximum value are higher for the group working with the knowledge text. Quartiles and medians for both groups are depicted by the box plots, see Figure 2. There are no outliers in the dataset of the variable t_1 for any group.

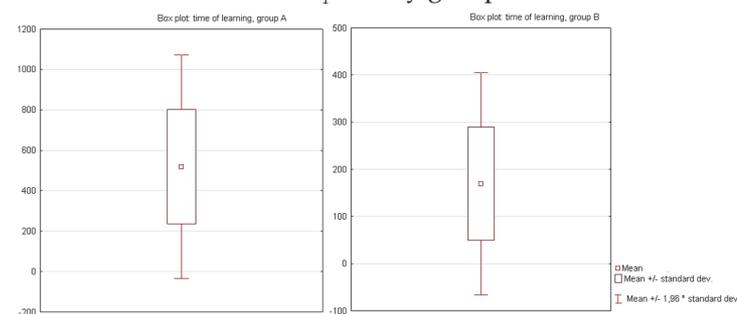


Figure 2: Box plots, variable t_1 , week 1

Distribution of frequencies - accuracy of the problems solved

We determine the distribution of frequencies for the variable acc_i for both groups A and B and i -th week. We use the bivalent scale, where

- “1” means “the problem was solved correctly”;
- “0” means “the problem was solved with an error”.

The statistics are summarized in Table 2.

Category	Group A		Group B	
	Frequency	Cumulative Frequency	Frequency	Cumulative Frequency
0	3	3	10	10
1	16	19	12	22

Table 2: Distribution of frequencies, variable acc_t , week 1

Group A reached a higher frequency of correct answers than the group B. The users working with the knowledge text achieved about 30% more accuracy than the users working with the standard text.

Testing of the statistical hypotheses - time of learning

For this purpose, we use parametric tests. Thus, we verify the normality of distributions for the variables tested in the experimental group A and the control group B. Shapiro-Wilk W test is used, see Table 3.

Null hypothesis H_0	The distribution of the basic set t_1 is normal	
Alternative hypothesis H_A	The distribution of the basic set t_1 is not normal	
Results for the group	p -value	α
Group A	0.82945	0.05
Group B	0.05287	0.05

Table 3: Normality of distribution test, variable: t_t , week 1

In any case, we cannot reject the null hypothesis; p -value $> \alpha$ for the groups A and B. It allows us to suppose that the assumption of normality of distribution is valid. Also the above-mentioned results can be confirmed visually, see the histograms in Figure 3.

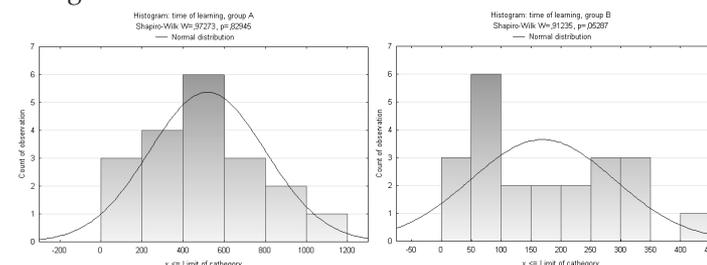


Figure 3: Histograms for the variable t_t , group A (left), group B (right), week 1

Two-sample tests of the significance of differences of sample means for t_1 between the groups A and B

Firstly we calculate two-sample F-test for the variance, see Table 4.

Null hypothesis H_0	The variance of the variable t_1 are equal ($\alpha = 0.05$)	
Alternative hypothesis H_A	The variance of the variable t_1 are not equal	
Statistics	p	a
Group A vs. Group B	0.00032	0.05

Table 4: Two-sample F-test for the variance, variable: t_t , week 1

As $p < \alpha$ on the level of significance $\alpha = 0.05$, we reject the null hypothesis H_0 . For further calculations we suppose that the variances of the data sets are significantly different.

In order to test the hypothesis on the equality of means we use the Behrens-Fisher test. The results are as follows.

Null hypothesis H_0	The means of the variable t_1 are equal ($\alpha = 0.05$)	
Alternative hypothesis H_A	The means of the variable t_1 are not equal	
Statistics	p	α
Group A vs. Group B	0.00005	0.05

Table 5: Equality of sample means test, variable: t_1 , week 1

As $p < \alpha$ on the level of significance $\alpha = 0.05$, we reject the null hypothesis H_0 . The statistically significant difference between sample means was confirmed.

Week 2 - Data Analysis

Descriptive statistics - time of learning

Also for the week 2 we provide basic descriptive statistics for the variable t_2 . In this week, the experimental group A was working with the standard text and the control group B was working with the knowledge text. The basic statistics are summarized in Table 6.

	n	Mean	Median	Min	Max	Variance	Standard deviation
Group A	19	393.68	296.00	7.00	1192.00	108614.10	329.57
Group B	22	412.00	383.50	62.00	935.00	62810.90	250.62

Table 6: Descriptive statistics, variable: t_2 , week 2

The mean value and the minimum value are higher for the group B working with the knowledge text, the variance and the maximum value are higher for the group A working with

the standard text. Quartiles and medians for both groups are depicted by the box plots, see Figure 4. There are no outliers in the dataset of the variable t_2 for any group.

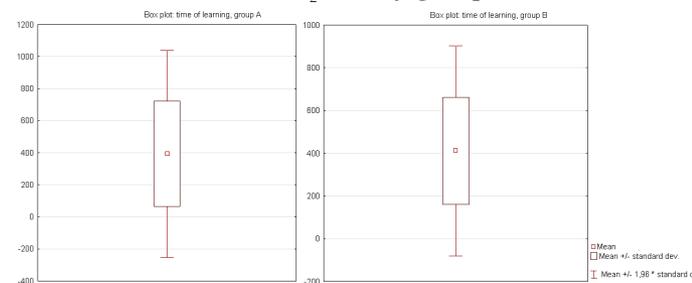


Figure 4: Box plots, variable t_2 , week 2

Distribution of frequencies - accuracy of the problems solved

We determine the distribution of frequencies for the variable acc_2 for both groups A and B. Again, the bivalent scale (1 - solved correctly; 0 - solved incorrectly) is used. The statistics are summarized in Table 7.

Category	Group A		Group B	
	Frequency	Cumulative Frequency	Frequency	Cumulative Frequency
0	6	6	6	6
1	13	19	16	22

Table 7: Distribution of frequencies, variable acc_2 , week 2

Group B reached a higher frequency of correct answers than the group A. The users working with the knowledge text achieved about 5% more accuracy than the users working with the standard text.

Testing of the statistical hypotheses - time of learning

We use the same approach as for the Week 1. To verify the normality of distributions for the variables tested in the experimental group A and the control group B, we use the Shapiro-Wilk W test, see Table 8.

Null hypothesis H_0	The distribution of the basic set t_2 is normal	
Alternative hypothesis H_A	The distribution of the basic set t_2 is not normal	
Results for the group	<i>p</i> -value	α
Group A	0.05772	0.05
Group B	0.15570	0.05

Table 8: Normality of distribution test, variable: t_2 , week 2

In any case, we cannot reject the null hypothesis; p -value $>$ α for the groups A and B. It allows us to suppose that the assumption of normality of distribution is valid. Also the above-mentioned results can be confirmed visually, see the histograms in Figure 5.

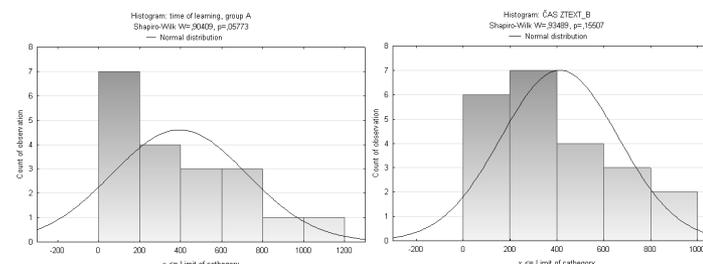


Figure 5: Histograms for the variable t_2 , group A (left), group B (right), week 2

Two-sample tests of the significance of differences of sample means for t_2 between the groups A and B

Firstly we calculate two-sample F-test for the variance, see Table 9.

Null hypothesis H_0	The variance of the variable t_2 are equal ($\alpha = 0.05$)	
Alternative hypothesis H_A	The variance of the variable t_2 are not equal	
Statistics	<i>p</i>	α
Group A vs. Group B	0.22981	0.05

Table 9: Two-sample F-test for the variance, variable: t_2 , week 2

As $p >$ α on the level of significance $\alpha = 0.05$, we cannot reject the null hypothesis H_0 . For further calculations we suppose that the variances of the data sets are not significantly different.

In order to test the hypothesis on the equality of means, we use the two-sample t test. The results are as follows.

Null hypothesis H_0	The means of the variable t_2 are equal ($\alpha=0.05$)	
Alternative hypothesis H_A	The means of the variable t_2 are not equal	
Statistics	p	α
Group A vs. Group B	0.841090	0.05

Table 10: Equality of sample means test, variable: t_2 , week 2

As $p > \alpha$ on the level of significance $\alpha = 0.05$, we cannot reject the null hypothesis H_0 . We cannot confirm that the difference between the sample means is statistically significant.

Dependence analysis of qualitative variables

We use the dependence analysis of qualitative variables to evaluate a feedback from the participants of the experiment. The feedback questions are as follows:

1. Subjective evaluation of the understandability of individual text styles
 - a) Both knowledge and standard texts are understandable for me.
 - b) Knowledge text is understandable for me, but the standard one is not.
 - c) Standard text is understandable for me, but the knowledge one is not.
 - d) Neither knowledge nor standard texts are understandable for me.
2. Prior knowledge of the algorithms tested
 - a) I was familiar with both the CPM method and the Dijkstra's algorithm.
 - b) I was familiar with the Dijkstra's algorithm only.

- c) I was familiar with the CPM method only.
- d) I was familiar neither with the CPM method nor the Dijkstra's algorithm.

Contingence tables and Chi-square test are being used for this purpose. The p -value allows us to confirm or reject the validity of the following hypotheses:

Hypothesis 1: There is no statistically significant dependence between a number of correctly-solved tasks and the prior knowledge of the algorithm tested.

We summarize the results of the experiment in Table 11.

Prior knowledge \ Accuracy	CPM and Dijkstra	Dijkstra only	CPM only	None of them
CPM and Dijkstra	6	0	3	1
Dijkstra	6	1	3	1
CPM	1	2	2	0
None of them	8	5	0	2

Table 11: Accuracy of the problem solving subject to the prior knowledge

We calculate the expected frequencies (see Table 12) and the p -value. As $p = 0.22235$, we cannot reject the Hypothesis 1 on the level of significance $\alpha = 0.05$ ($p > \alpha$). There is no statistically significant dependence between the prior knowledge of the algorithms and the accuracy of the problem solving.

Accuracy \ Prior knowledge	CPM and Dijkstra	Dijkstra only	CPM only	None of them
CPM and Dijkstra	5.12	1.95	1.95	0.98
Dijkstra	5.63	2.15	2.15	1.07
CPM	2.56	0.97	0.98	0.49
None of them	7.68	2.93	2.93	1.46

Table 12: Accuracy of the problem solving subject to the prior knowledge - expected frequencies

Hypothesis 2: There is no statistically significant dependence between correctly-solved tasks and a subjective evaluation of the understandability of the texts.

We summarize the results of the experiment in Table 13.

Accuracy \ Positive understandability	CPM and Dijkstra	Dijkstra only	CPM only	None of them
Both standard and knowledge texts	6	2	2	0
Knowledge text only	6	1	2	3
Standard text only	5	1	3	1
None of them	4	4	1	0

Table 13: Accuracy of the problem solving subject to the understandability of the texts

We calculate the expected frequencies (see Table 14) and the p -value. As $p = 0.351211$, we cannot reject the Hypothesis 2 on the level of significance $\alpha = 0.05$ ($p > \alpha$). There is no statistically significant dependence between the understandability of the texts and the accuracy of the problem solving.

Prior knowledge \ Positive understandability	CPM and Dijkstra	Dijkstra only	CPM only	None of them
Both standard and knowledge texts	5.12	1.95	1.95	0.98
Knowledge text only	6.15	2.34	2.34	1.76
Standard text only	5.12	1.95	1.95	0.98
None of them	4.61	1.76	1.76	0.88

Table 14: Prior knowledge subject to understandability – expected frequencies

Hypothesis 3: There is no statistically significant dependence between the prior knowledge of algorithms tested and the subjective evaluation of the understandability of the texts.

We summarize the results of the experiment in Table 15.

Prior knowledge \ Positive understandability	CPM and Dijkstra	Dijkstra	CPM	None of them
Both standard and knowledge texts	3	3	4	0
Knowledge text only	1	4	3	4
Standard text only	2	1	5	1
None of them	4	3	3	0

Table 15: Prior knowledge subject to the understandability

We calculate the expected frequencies (see Table 16) and the p -value. As $p = 0.223228$, we cannot reject the Hypothesis 3 on the level of significance $\alpha = 0.05$ ($p > \alpha$). There is no statistically significant dependence between the understandability of the texts and the prior knowledge of the algorithms.

Accuracy Positive understandability	CPM and Dijkstra	Dijkstra	CPM	None of them
Both standard and knowledge texts	2.44	2.68	3.66	1.22
Knowledge text only	2.93	3.22	4.39	1.46
Standard text only	2.20	2.41	3.29	1.10
None of them	2.44	2.68	3.66	1.22

Table 16: Accuracy of the problem solving subject to the understandability- expected frequencies

Hypothesis 4: There is no dependence between the type of the texts and the accuracy of the problem solving

We summarize the results of the experiment in Table 17.

Problem solved Type of the text	Week 1		Week 2	
	Correctly	With errors	Correctly	With errors
Knowledge text	16	3	16	6
Standard text	12	10	13	6

Table 17: Accuracy of the problem solving subject to the type of the texts

We calculate the expected frequencies (see Table 18) and the p -values for both weeks. In the Week 1, as $p = 0.041799$, we reject the Hypothesis 4 on the level of significance $\alpha = 0.05$ ($p < \alpha$). There is statistically significant dependence between the accuracy of the problem solved and the type of the text. In the Week 2, as $p = 0.762504$, we cannot reject the Hypothesis 4 on the level of significance $\alpha = 0.05$ ($p > \alpha$). There is no statistically significant dependence between the accuracy of the problem solved and the type of the text.

Problem solved Type of the text	Week 1		Week 2	
	Correctly	With errors	Correctly	With errors
Knowledge text	12.98	6.02	15.56	6.44
Standard text	15.02	6.98	13.44	5.56

Table 18: Accuracy of the problem solving subject to the type of the texts - expected frequencies

Summary of the statistical analysis

We provide the summary of the statistical analysis separately for individual weeks of the experiment and for the complete experiment.

Week 1

Mean values as well as variances for the variable “time of learning” are significantly different between the experimental group and the control group. There is statistically significant dependence between the type of the texts and the accuracy of the problem solving.

Week 2

Neither mean values nor variances for the variable “time of learning” are significantly different between the experimental group and the control group. Also there is no statistically significant dependence between the type of the texts and the accuracy of the problem solving.

Complete experiment

Based on the analysis of the qualitative variables provided by the participants of the experiment as the feedback from them, we can conclude that

- there is no statistically significant difference between a number of the correctly-solved tasks and the prior knowledge of the algorithms tested;
- there is no statistically significant difference between a number of the correctly-solved tasks and the subjective evaluation of the understandability of knowledge and standard texts;
- there is no statistically significant difference between the subjective evaluation of the understandability of knowledge and standard texts and the prior knowledge of the algorithms tested.

Discussion

After we have compared our approach and results achieved with the approaches and results by other authors, we can explain the reason of our findings, where only one partial characteristic was found as statistically significant. Following Peng and Hengartner (2002), the main proof lies in the literary styles of texts used in our experiment. Literary style is an important characteristic of a text and can be measured objectively using statistical methods to distinguish styles of individual authors. The technique can be even use to recognize the author of an anonymous text (Wan et al., 2012).

Although Peng and Hengartner applied their method to the texts of classic literature (by Shakespeare, Dickens, etc.), also their approach would be helpful to measure the similarity between the standard and knowledge text styles. We assume that subject to very formal structure of any text describing mathematical methods (the CPM method and Dijkstra algorithm, in our case), there are no statistically significant differences between the two styles of text used in our experiment.

This point also could elucidate why our results do not correspond with the findings by Ozuru et al. (2009). They showed the impact of prior knowledge on reading skills and text comprehension. Perspectively, we could use their methodology in a reverse way; not to measure the rate of influencing the text comprehension through the prior knowledge, but to filter the impact of it and actually eliminate the initial differences in the prior knowledge among the participants of the experiment. As Tarchi (2010) showed, similar experiment including statistical analysis (multiple linear regression analysis) could be processed correctly even for such an unstructured branch as the history is.

Conclusion

In this paper we present a methodology how to carry out an experiment to verify the impact of the style of an educational text on learning outcomes and students' performance. We used the methods of statistical analysis and accept or reject several hypotheses formulated. Even though only one partial hypothesis can be accepted, the research opens many new ways to improve the experiment.

In Materials and Methods, we presented a methodology to create the knowledge text from the standard one. In the future work we feel necessary to develop the methodology and the procedure of the knowledge text creation in more details. Also the application domain for the experiment should be selected more carefully; we change our focus on less formalized areas than applied mathematics or operations research are.

The main topic for the future work is to select measures for establishing the metrics to quantify the similarity of standard and knowledge texts covering the same contents, which is

presented by different ways only. This allows us to construct the texts of really different styles to measure the impact of such styles to key variables of the learning outcomes.

Acknowledgements

The paper is supported by the grant project of the Internal Grant Agency of the Czech University of Life Sciences Prague "Measuring the efficiency of knowledge transfer in the agricultural waste processing sector", No. 20131001.

References

- Dobrylovský, J. (2009) 'Dlouhodobé trendy ve vývoji obsahu učebnic zeměpisu pro gymnázia', *E-Pedagogium*, vol. 2009, no. 5, pp. 7-25.
- Dömeová, L., Houška, M. and Beránková, M. (2008) *Systems Approach to Knowledge Modelling*, Hradec Králové: GSOC.
- Glava, A. and Glava, C. (2011) 'The model and the didactic modelling an analytic point of view', *Procedia - Social and Behavioral Sciences*, vol. 15, pp. 2228-2231.
- Hodis, V. (2003) Hodnocení učebnice Praktické činnosti pro 6. – 9. ročník základních škol: Práce s počítačem. In: *Teoretické i praktické problémy edukaci technické i informatické 2003*. Rzeszów: Wydawnictwo Oświatowe FOSZE, pp. 143-147.
- Kools, M., Van De Wiel, M.W.J., Ruiter, R.A.C., Crüts, A. and Kok, G. (2006) 'The effect of graphic organizers on subjective and objective comprehension of a health education text', *Health Education and Behavior*, vol. 33, no. 6, pp. 760-772.
- Kolman, L. (2008) 'K-Maps of HR Selection Procedures and Their Potential Use in the Identification of Talented Students', *Journal on Efficiency and Responsibility in Education and Science*, vol. 1, no. 1, pp. 20-26.

- Lee, J.H. and Segev, A. (2012) 'Knowledge maps for e-learning', *Computers and Education*, vol. 59, no. 2, pp. 353-364.
- Lindseuy, J. K. (2009) *Introduction to Applied Statistics: A Modelling Approach*. 2nd Edn. New York: OXFORD University Press.
- Maňák, J. and Janík, T. (2009) Výukové metody jako předmět výzkumu. In: M. Janíková and K. Vlčková, eds.: *Výzkum výuky: tematické oblasti, výzkumné přístupy a metody*. Brno: Paido, pp. 83-96.
- Mikk, J. (2007). Učebnice: budoucnost národa. In: Maňák, J. and Knecht, P. eds. *Hodnocení učebnic*, Brno: Paido, pp. 11-23.
- Ozuru, Y., Dempsey, K. and McNamara, D. S. (2009) 'Prior knowledge, reading skill, and text cohesion in the comprehension of science texts', *Learning and Instruction*, vol. 19, no. 3, pp. 228-242.
- Peck, R. and Devore, J. L. (2012) *Statistics: The Exploration and Analysis of Data*. 7th Edn. Boston: Brooks/Cole Cengage Learning.
- Peng, R.D. and Hengartner, N.W. (2002) 'Quantitative analysis of literary styles', *The American Statistician*, vol. 56, no. 3, pp. 175-185.
- Prasad, G.V.R.J.S. and Ojha, A. (2012) 'Text, table and graph - Which is faster and more accurate to understand?', *Proceedings - 2012 IEEE 4th International Conference on Technology for Education, T4E 2012*, pp. 126-131.
- Starý, K. and Chvál, M. (2009) Kvalita a efektivita výuky: metodologické přístupy. In: M. Janíková and K. Vlčková, eds.: *Výzkum výuky: tematické oblasti, výzkumné přístupy a metody*. Brno: Paido, pp. 63-81.
- Šubrt, T. et al. (2011) *Ekonomicko-matematické metody*, Plzeň: Aleš Čeněk.

Tarchi, Ch. (2010) 'Reading comprehension of informative texts in secondary school: A focus on direct and indirect effects of reader's prior knowledge', *Learning and Individual Differences*, vol. 20, no. 5, pp. 415-420.

Tannenbergová, P. (2009) 'Učebnice dějepisu pro 6. ročník základní školy z pohledu jejich didaktické vybavenosti', *E-pedagogium*, vol. 2009, no. 5, pp. 108-121.

Tudor, S. L. (2012) 'A study on the efficiency of using combined modern and traditional didactic strategies', *Procedia - Social and Behavioral Sciences*, vol. 33, pp. 989-993.

Vališová, A. and Kasíková, H. (2011) *Pedagogika pro učitele*, Prague: Grada Publishing.

Wan, Ch. H.; Lee, L. H., Rajkumar, R. and Isa, D. (2012) 'A hybrid text classification approach with low dependency on parameter by integrating K-nearest neighbor and support vector machine', *Expert Systems with Applications*, vol. 39, no. 15, pp. 11880-11888.

Zack, M. H. (1999) 'Managing Codified Knowledge', *Sloan Management Review*, vol. 40, no. 4, pp. 45-58.