Large scale studies with concept mapping

Estudios a gran escala con mapas conceptuales

Katrin Soika
Tallinn University, Estonia

Priit Reiska
Tallinn University, Estonia

Large scale studies with concept mapping

Estudios a gran escala con mapas conceptuales

Katrin Soila, Tallinn University, Estonia
katrin.soika@gag.ee

Priit Reiska, Tallinn University, Estonia
priit@tlu.ee

Abstract
Concept mapping is a method for determining the achievement of knowledge. Concepts are linked with labelled lines to proposition and so the concepts create a graphical structured meaningful relationship. There are many ways to use concept mapping in research as data collecting and assessment instrument. Changing the conditions (like focus question about the concept map, lists of concepts, given structure of concept map etc.) also change the results. For a valid research is necessary to analyse the study and define the aims before collecting the data. Probably the most comfortable concept mapping constructing opportunity is to use special Internet based environment and analysing program – that makes data collection easier and more objective. This article brings out, what kind of problems may occur, when concept mapping method is used as a research instrument in a large scale study and it also tries to define how to select the a valid instrument for a study. Researchers want to analyse students' knowledge, but instead sometimes they can only control whether they were able to create concept maps. The study brought out, that the quality of concept maps does not depend on concept maps creating frequency and computer handling skills.

Resumen
Los mapas conceptuales son un método para determinar los logros de aprendizaje. Los conceptos se unen mediante líneas etiquetadas y, de esta forma, los conceptos crean una relación significativa estructurada gráficamente. Existen muchas vías para usar mapas conceptuales en investigación como instrumento de recogida de datos y cómo instrumento de evaluación. Cambiando las condiciones (como la pregunta de enfoque, la lista de conceptos, una estructura dada del mapa conceptual, etc.) también cambian los resultados. Probablemente la forma más cómoda de construir mapas conceptuales sea usar un programa de análisis basado en un entorno de Internet, que hace la toma de datos más fácil y objetiva. Este artículo resalta los distintos problemas que pueden aparecer cuando los mapas conceptuales son usados como instrumento de investigación en un estudio a gran escala e intenta definir también cómo seleccionar un instrumento válido para un estudio. Los investigadores desean analizar el conocimiento de los estudiantes, pero en muchos casos sólo controlan si fueron capaces de crear mapas conceptuales. El estudio hizo evidente que la calidad de los mapas conceptuales no depende de la frecuencia de realización de mapas conceptuales ni de las habilidades para el manejo del ordenador.

Keywords
Concept mapping, assessment, validity, large scale study

Palabras clave
Mapas conceptuales, evaluación, validez, estudio a gran escala

1. Introduction and theoretical background

This article aims to find out what are the main problems which may occur when concept mapping method is used as a research instrument for large scale studies. In this article all of the concept maps and analyses of results are made using a computer programme. It is explained why it is comfortable and objective. This article describes how to control researching instruments' validity.

Concept mapping method

Concept mapping method is developed by Joseph Novak and his research team in early 1970s (Novak, 2010). The method is based on the theory of Ausubel (1968). This is also called meaningful learning and it assumes that learners construct their knowledge while they are already influenced from the previous knowledge. Concept maps could be created in different ways – on computer, by pen and paper, with labels etc. (Reiska et al., 2008; Ruiz-Primoet al., 1997; Novak, 2010). Concept map is a collection of propositions which is constructed in a certain way; it expresses graphically structured meaningful relationships which exist between different concepts (Ruiz-Primo et al., 1997). Novak (2010) suggests connecting the concepts with lines in order to create a good concept map; and labelling the lines with one or a few linking words which define the relationship between the two concepts so, that one could read them as a statement or a proposition.

Concept maps are widely used at schools in learning process; the method helps to prevent rote learning, to summarize already studied knowledge or class discussions, to create presentations etc. (Novak, 2010). Educational mapping is seen as a tool, which supports activities of learning, teaching, researching etc. The main idea in all fields of concept mapping is to reflect the brain work. Educators are interested in connections of working and long-term memory.

The structure of the concept map depends on many different conditions. Before students are asked to create a map, it should be clarified what kind of knowledge the map should develop or assess. There are many different possibilities to instruct students. Concept maps could be created without conditions, with a focus question, with root concept, with a list of concepts, restricting list of concepts, expert skeleton concept maps etc. Each of the methods gives a different map, with various outlook and nature (Cañas et al., 2012). To use concept mapping for assessment should be clarified the validity of the concept mapping. It should be verified concept maps are checking the knowledge of students, not constructing ability, computer handling skills or something else. For example if to provide students only with a focus question, it cannot be checked how all the students use some concept, which is connected with the question.

Studies from science have proved the necessity of concept maps in assessing and have pointed out that concept mapping is a useful tool to portrait the process of knowledge transformation from novice to expert. At first, concept maps were created mainly by pen and paper and results were calculated manually. The whole process took quite long.

Although many researchers have reported that concept mapping is a useful tool for learning and instruction, scientists have found some disadvantages to constructing concept maps using pencil and paper:

a) It is inconvenient for a teacher to provide appropriate feedback to students during concept mapping.
b) The construction of a concept map is complex and difficult for students, especially novice students.
c) Concept maps constructed using pencil and paper are difficult to revise. The 'pencil-and-paper' concept map is not an efficient tool for evaluation. (Chang et al., 2005).

Nowadays it has become easier because of the computers. (Novak, 2010; Gouli et al., 2003).

Assessing with concept maps

Novak (2010) noticed that students had recognized the value of concept mapping as learning and assessing tool.

For evaluating concept maps, certain dimensions for measuring are needed. Miller, Cañas et al (2008) have developed a topological taxonomy for evaluating created concept maps. They

considered five criteria, when topological levels were defined: 1) recognition and using concepts; 2) presence of linking phrases; 3) degree of ramification; 4) hierarchical depth and 5) presence of cross-links. The taxonomy consists of 7 levels: from 0 to 6. Maps, which are evaluated with 5 and 6, satisfied almost all of the criteria. There are several measures for analysing concept maps: number and quality of propositions, size and hierarchy of the concept map, clusters of maps (Reiska et al., 2008).

An analyzing program is created for a programme IHMCmaps (http://cmap.ihmc.us/)- *CmapAnalysis.* This program gives an opportunity to analyse various algorithms, rubrics and techniques of concept maps. Parameters could be defined by the researcher. Creators of the program propose that this software helps instructors, researchers and teachers to have routine analytical operations automatically (Cañas et al., 2010). This program is indispensable for assessing and analyzing concept maps of large scale studies.
Sometimes it is said that concept maps could be used for assessing only when they have also been used in the learning process. This argument, however, is debateable and therefore further analysed in the current study.

2. Research questions and methodology

The main task of the paper was to study the validity of using concept mapping for assessment. More specifically, to find out what are the issues assessed using concept mapping. For that, the following research questions were defined:

a) Why are concept maps marked with different taxonomy scores? Could the computer based analysing program help in understanding the map quality?
b) Does the concept mapping quality measure, like taxonomy score and number of propositions, depend on students knowledge or methodical skills (earlier experience uisng this method, easiness to use, etc)?

The data collection was carried out in 2012 and the stratified sample of the study included 1614 sixteen to seventeen-year old students from 46 Estonian high schools. Schools varied by location, number of students, results of the state exams. The aim of the study was to analyse natural science literacy skills among Estonian students. Data collection instruments were: PISA-like three dimensional scenario-based tasks from chemistry, physics, geography and biology; concept map from one of the subjects and questionnaires about previous use of concept mapping and students' computer skills. Some of the questions were open ended and some were with multiply choices.

This paper is focused on 377 students, who created a concept map in the field of chemistry. Students had a focus question and a list with 30 certain concepts, meaning they did not have much freedom. The maps and answers to the questionnaire were analysed for this paper.

Data collection was performed with different instructors who used the Internet version of programme *CmapTools* and the same instructions were given for the whole class. Students needed computers and Internet connection for the mapping. Each student had a personal password and code for the Internet based *CmapTools*.

Problems with large scale study data collection and analysis

The main problem with large scale study data collection and analysis is how to carry out concept mapping so that it would measure the knowledge, not the ability to create concept maps or to use the computer. The pilot study compared what kind of concept maps are created by the students, when: a) only a focus question and the solved scenario-based exercise are given to students; b) focus question and the list on concepts (defined by the scenario-based exercise) are given to students. The result was that highly taxonomy scored concept maps were built when students had a focus question and the list of concepts (Soika et. al., 2012). This pilot study showed that it is difficult to measure with concept mapping the same aspects (e.g. conceptual knowledge) if just one condition is different.

In current study the students had to solve the scenario-based exercise on the first day. Weeks later they had to create the concept map about the same topic. Some of the data was lost, because some students were absent. There was also some data lost because of the internet based concept mapping environment (all students could not save their work; school did not have the right hardware

in all the computers; some computers broke suddenly; Wifi connection was not strong enough). In order to eliminate such problems to minimum, the schools were checked before. It would have been easier with less students and schools. Internet based environment was used instead of computer based program, because it was more comfortable and a special code was given to every student that did not strain schools servers. Nobody was therefore able to peek into others work.

The questionnaire was created in Google Docs environment; some students forgot to submit their answers. For eliminating such problem, students were asked several times to follow the instructions (instructions were in written form).

Students were asked to compose their maps individually and different topics were distributed all over the classroom. Still students wanted to collaborate, to use help from the internet database or to connect concepts randomly without delving. Some of the students were a bit afraid of assessing and wanted to show that they are able to connect all of the concepts. Instructors had said that the aim of the work was not to connect the concepts randomly, but to create correct propositions. In order to eliminate such problems to minimum, all of the instructors had been instructed and used the same presentations. This also gave a possibility to get comparable concept maps and validity results for the research.

Some more problems arouse when analysing concept maps with the programme due to the peculiarities of the Estonian language - some special letters (õ, ö, ä, ü) are not accepted by the analysing programme. Therefore some words from fails had to be rewritten. Due to the large number of concept maps, this process was quite long-lasting.

Whether it is possible to compare the current study with other similar concept maps based studies is questionable. As the conditions for compiling the concept maps are not always known, the results cannot be compared. But there is always a possibility to glean some important facts to the instruments and studies from other interesting concept mapping based studies. In large scale studies with concept maps which have many different instructors have to be very careful with establishing the same conditions every time, otherwise the results would not be comparable.

Conditions for the concept mapping task

The focus question of the concept map was: "Cold bag- is it only chemistry?". The question was connected to previously solved exercise. Students received 30 different concepts, which were defined by experts. Concepts were in different abstract levels and from various subjects and topics of natural sciences. 10 concepts were from chemistry (water, solubility, exothermic reaction, endothermic reaction, speed of reaction, equilibrium of chemical reaction, mole, pH, temperature of freezing, water, salt), 6 from physics (energy transfer, energy, pressure, melting, friction, absorption), 6 from biology (capillary, nerve impulse, lymphatic drainage, blood circulation, edema, dislocation) and 8 were social concepts (cold bag, tumour, risk, safety, pain, ethics, treatment, first aid). Time for creating a concept map was not limited.

3. Results of the study

Why are concept maps marked with different taxonomy scores? Could the computer based analysing program help us in understanding the map quality?

The average taxonomy score for all concept maps in the study was 2,5. There were two certain types of concept maps used and others were unclassifiable.
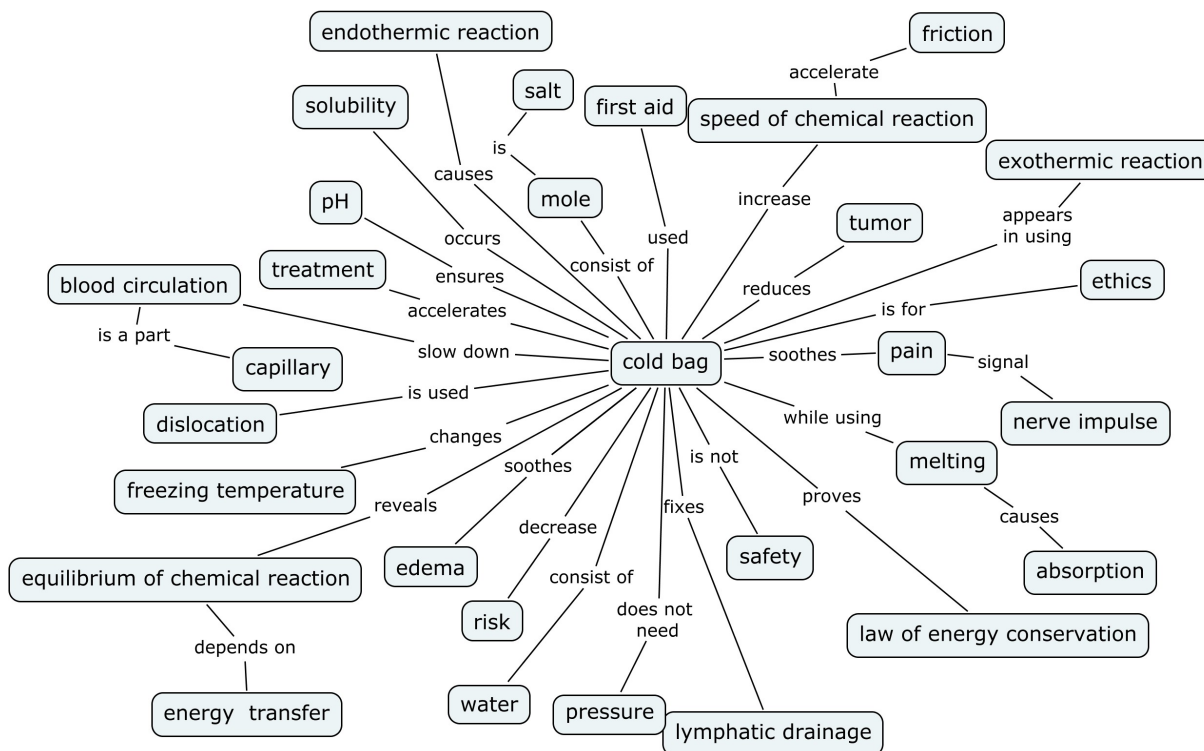
Cold bag- is it only a chemistry?

Figure 1. "Star"-shape concept map, which taxonomy score was 2 and average proposition per concept was 1, there was no orphan count, but the proposition count for concept cold bag was 24.

The concept map of 2% of the students who created 23-15 propositions with the concept cold bag was mainly with a shape of a "huge star". Concept "cold bag" was in the middle and all of the other concepts were connected mainly wit the central concept. Average taxonomy score for such maps was 2,25, that is lower than the average taxonomy score. In Figure 1 the map is not with high taxonomy scored, because students did not find connections with concepts to each other (for example in the concept map of Figure 1 the student could make a proposition "tumour could be pain"). So the score of taxonomy score does not depend only from the number of average proposition count, but it also depends on the structure of the whole map. Higher taxonomy scores for the maps groups appeared, when fewer concepts were connected with the central concept. These maps where only some propositions existed were marked with 1 or 0.

Table 1

Average proposition count per concept for the group	Average taxo-nomy score for the group	Number of proposition with the concept **cold bag**	% of students	Average proposition count per concept for the group	Average taxo-nomy score for the group	Number of proposition with the concept **cold bag**	% of students
0,77	2,25	17,9 *	2,1	0,78	2,76	6	9,5
0,73	2,36	14,7 **	2,9	0,69	2,54	5	9,3
0,93	3,14	10	3,7	0,61	2,71	4	13
0,88	3,09	9	2,9	0,57	2,55	3	15
0,72	2,67	8	4,8	0,56	2,56	2	16
0,78	2,56	7	4,8	0,44	1,78	1	9,8
				0,49	1,89	0	4,8

*- the concept "cold bag" was connected with 23-15 propositions- the number is average of the proposition count

**- the concept "cold bag" was connected with 14-11 propositions- the number is average of the proposition count
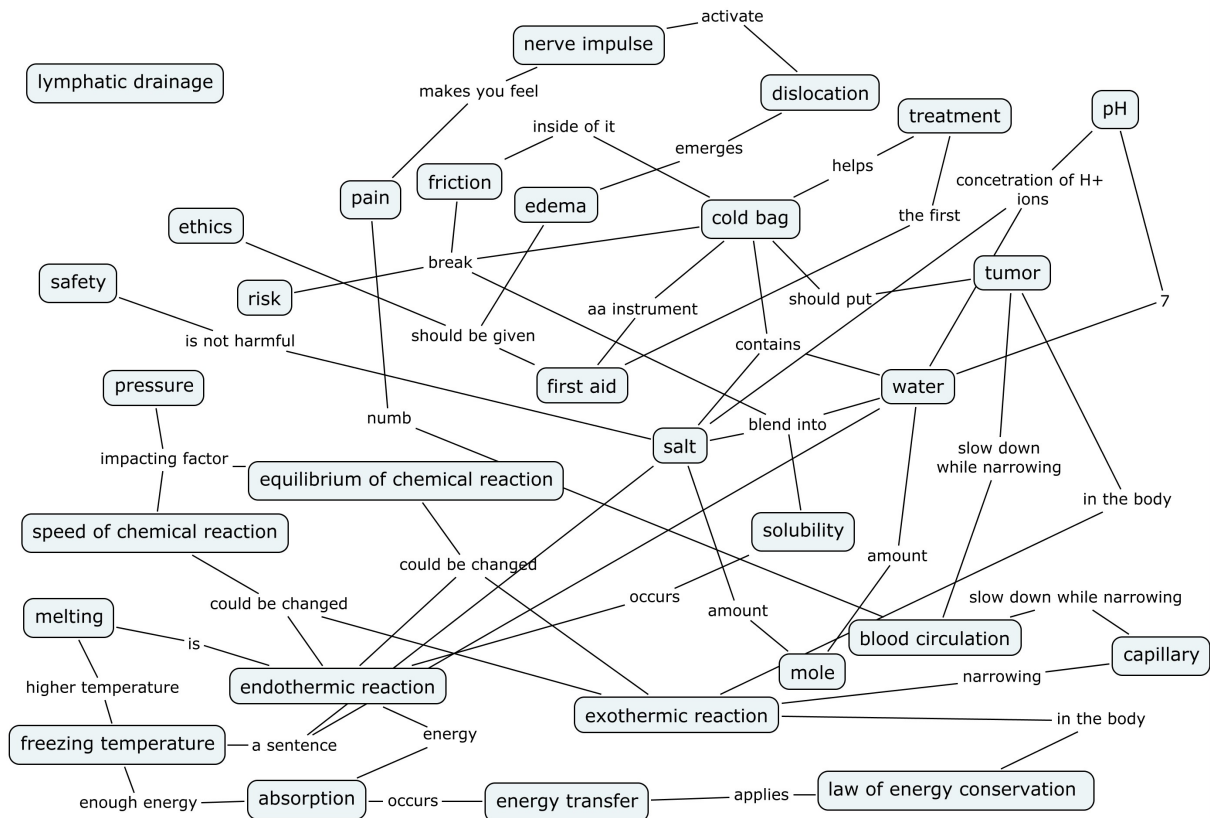
Cold bag- is it only a chemistry?

Figure 2. The taxonomy score of the map was 4 and the average proposition count was 1,5; there was 1 orphan count and the number of proposition for the concept cold bag was 10

Thanks to the computers and analysing programme, it is possible to imagine the shape of the map, when seeing only the scores of different indicators (as taxonomy score, central concept, orphan count etc.). That analysing method makes it easier to generalize the data. It is very hard to handle and analyse large scale studies of concept maps without the special program. At first, collecting the data would be hard and secondly, the analysis would take much time. The program enables higher level of objectivity, because the matrix measures the same values.

Does the concept mapping quality measure, like taxonomy score and number of propositions, depend on students knowledge or methodical skills (earlier experience uisng this method, easiness to use, etc)?

If the taxonomy scores depended on the program handling skills or the frequency of creating concept maps, the instrument of the study would not be validated, because the knowledge from the topic of students would not be checked. To exclude such possibilities, the questionnaire with multiply choices was used. The questionnaire included questions like how many times students had created concept maps before; how many times they had used computers for creating the maps; if they enjoyed creating concept maps and if they had some problems with creating the concept map during the study.

Analysis of the data revealed that there is no correlation between taxonomy score level, frequency of creating concept maps, program handling skills, pleasantness, school type or the number of classmates.

As table 2 shows, most of students from the sample had made concept maps before.

Table 2

How many times have you create concept maps?	More than 10	6-10	2-5	1	I have never made any map
% of students	30 %	17 %	38 %	3,7%	11%

A question arose whether the quality of concept map depends on the frequency of creating concept maps. This analyse bases on the average taxonomy scores.
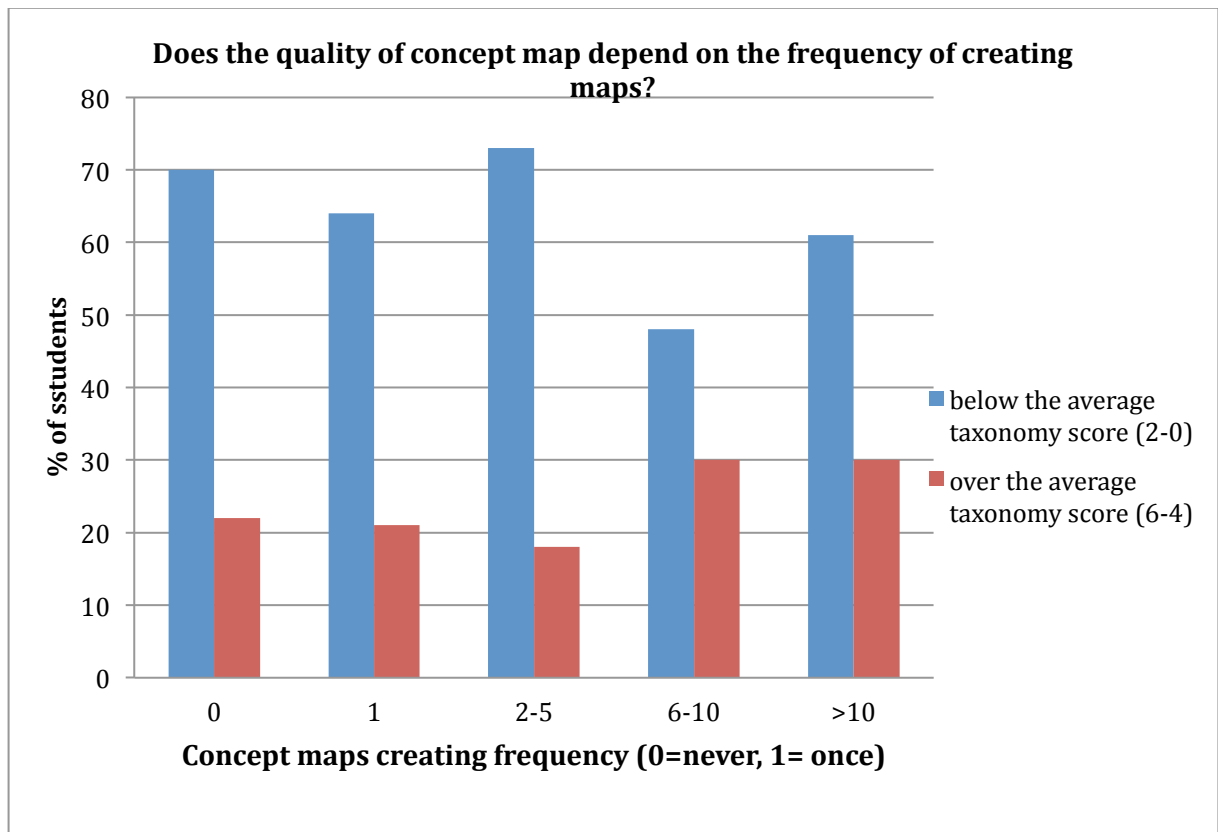

Figure 3. How does the quality of concept maps depend on the frequency of creating maps-comparing taxonomy score and concept maps creating frequency?

Figure 3 shows that the lowest taxonomy scores got maps that were created by students who had made concept maps 2-5 times. Students, who had never created concept maps before built more maps that scored 5 than students, who had made maps 6 to 10 times and they had more over the average taxonomy scored maps, than students, who had built maps 2 to 5 times.
It is also interesting whether the average proposition count depended on the concept maps creating frequency.
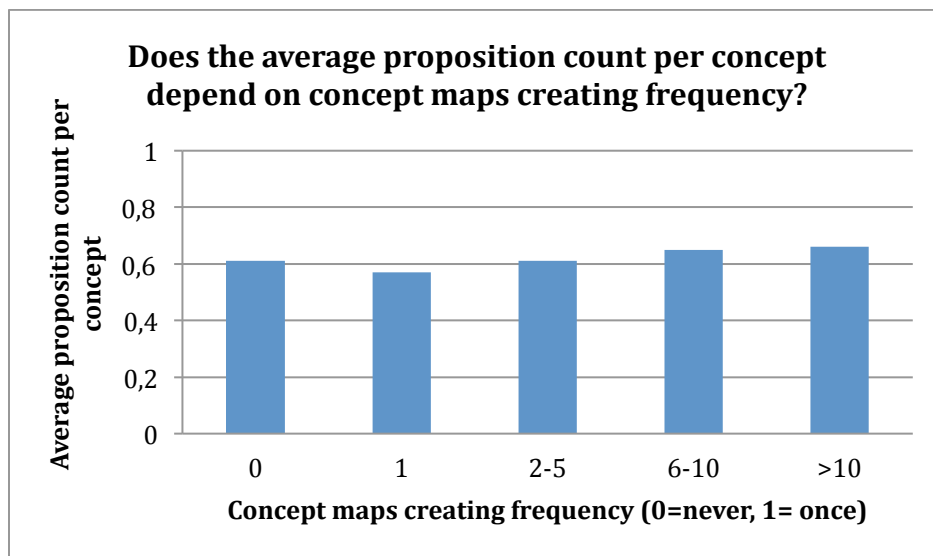
Figure 4. How does the quality of concept maps depend on the frequency of creating maps-comparing average proposition count and concept maps creating frequency?

Figure 4 reveales, that there is no huge difference in the number of average proposition count per concepts, when comparing the frequency of creating concept maps. Students, who had never made any concept maps, had as many propositions per concepts as students, who had created maps 2 to 5 times.

In the instrument, there were 9 abstract chemistry and physics-based concepts. Testing whether these concepts were used more by students, who had created more concept maps before, was carried out. An assumption was that probably students, who had created more concept maps before, have more time and skills to use abstract concepts in the propositions.
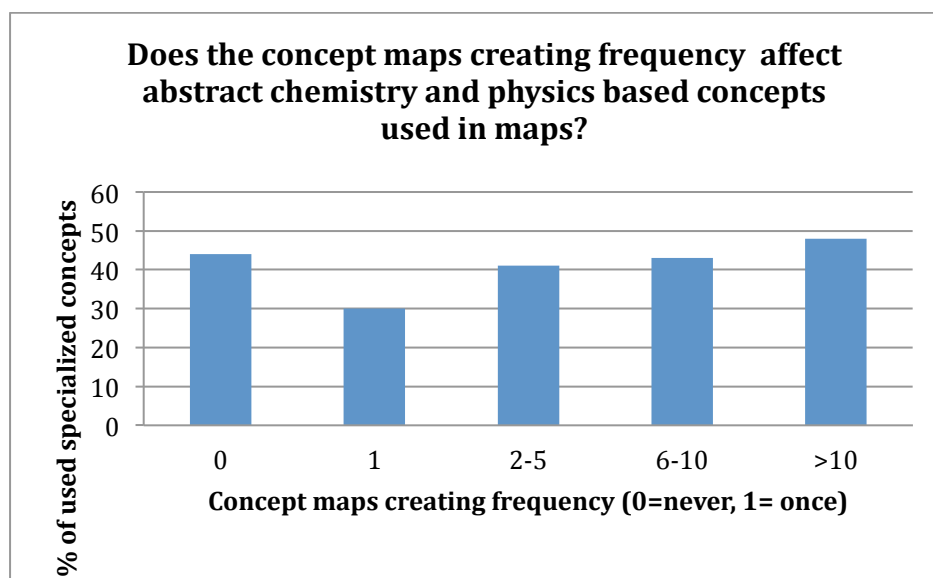

Figure 5. Do experienced concept mappers use more chemistry and physics based concepts than beginners?

Figure 5 shows that there is no differences between novices and expert students in using abstract subject based concepts. For the analysis, nine low centrality abstract concepts from the given concepts list were chosen. As seen from Figure 5, 44% of students, who had never created concept map, used all of the abstract concepts in their concept maps. 48% of the students, who had created concept map more than 10 times, also used all of the concepts.

So based on this study, it can be concluded, that the quality of the concept map does not depend on the concept map creating frequency and it was valid to use concept map in our study. Maybe the

computer-based program used in the study was not valid, because the computer handling skills were not measured. While analysing the data, a question arose: were the program handling skills measured? Figure 6 illustrates this question.
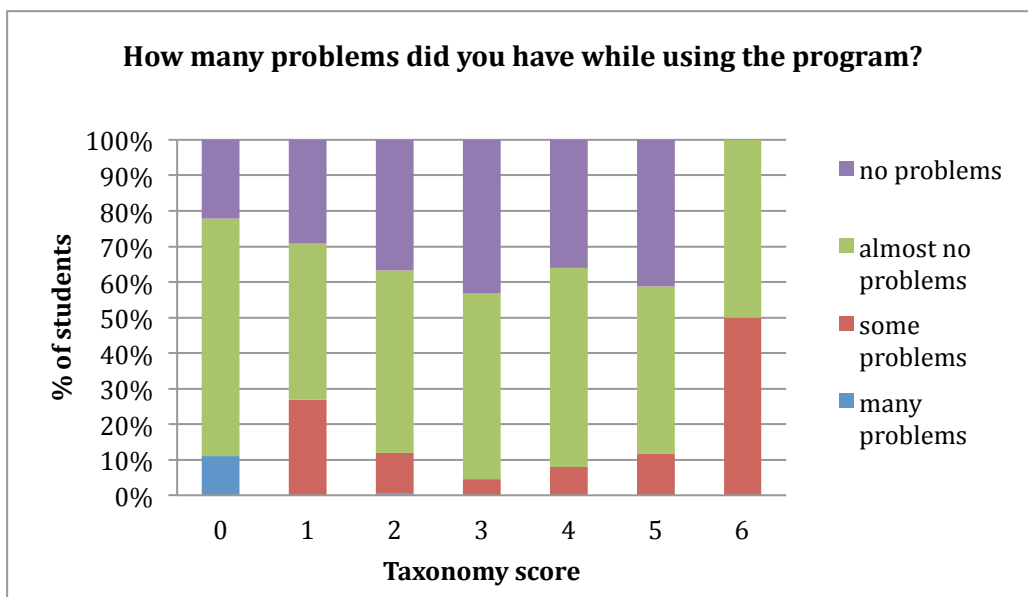

Figure 6. Does the taxonomy score of concept maps depend on problem appearance frequency?

As we seen from the Figure 6, most of the students said that they had almost no problems while they created their concept map in the computer. It is surprising, that students, whose concept maps got taxonomy score 6, found that they had some problems or almost no problems during the process. Therefore, students, whose map was evaluated with low level taxonomy score, got the mark due to their knowledge, not because of the occurred problems with program. Only 11% of students, who got taxonomy score 0, admitted that they had technical problems with the program (some of them had added, that they could not save or did not have the opportunity to concentrate, because the computer or the problematic Wifi connection). Therefore the program using skills were not measured and the instrument was valid.

It was also the interest of the researchers to find out who had more problems with the concept maps creating program- those, who had made concept maps often or those, who had not made concept maps.
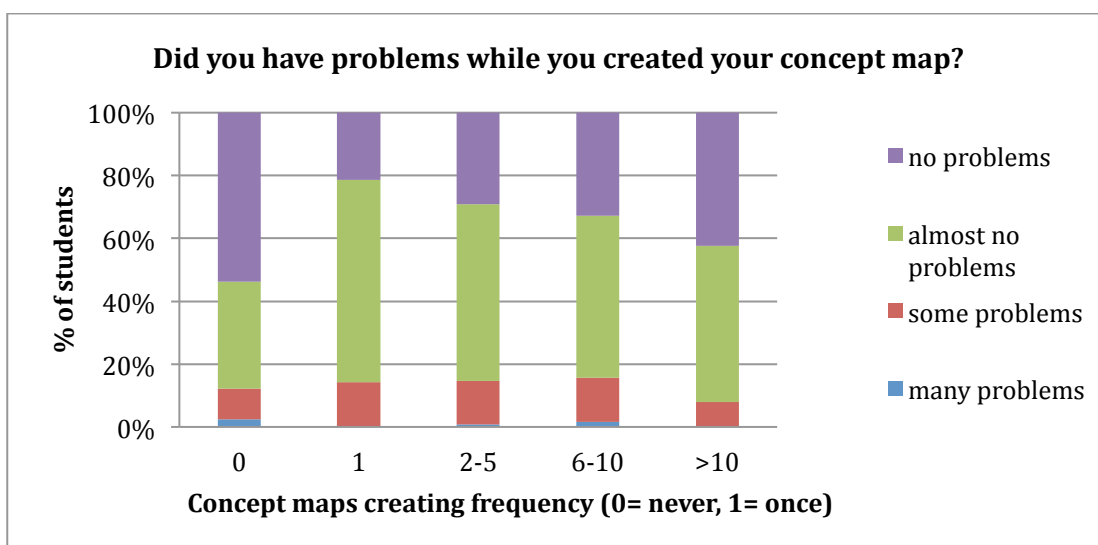

Figure 7. Does the problem appearance frequency depend on the concept maps creating frequency?

Figure 7 pointed out, that problems of the concept map building did not depend on the frequency of creating maps. Students, who had not created concept maps, did not have more problems with the program. Sometimes it is questioned whether it is correct to compare maps of students who are novices and experts in building concept maps. The current study revealed that in the internet-based environment, there was no difference. It means that the instructors had explained the topic clearly and so the students felt that they were able to fulfil this exercise- they were able to create a concept map with the focus question and the list of concepts. Therefore it could be concluded that the study was valid and the difference of computer handling or instructors competences was not measured.

Because of the comprehensive sample, the researchers decided to find out if the result of the concept map depended on the school feature. No correlation was found between schools based on its peculiarity (number of students, location, results of the state exams or directions) and taxonomy score results. During this study, the best concept maps were not created in bigger schools in (with more than 100 students in one grade level), but in the average and smaller ones.

4. Discussion

It is always hard to carry out a large-scale study. Probably it is especially hard with a concept mapping instrument. Usually the main problem is how to create equal (even the same) conditions to the whole group of participants. Without the same conditions, the results could not be compared and the study would not be valid. Instructors should give the same advice and actually they should even use the same words and examples, because concept map is a projection from ones knowledge. The knowledge depends on the memory- how many facts one remembers about different concepts, what kind of feelings one has. Even different words from various instructors may affect students' consciousness diversely. Using computers makes data collection and analysis easier. It also means that one could be more objective in interpreting the results and data collection conditions are similar. It should be remembered that slightly different studies could not be compared, if all of the conditions are not known.

In this article, the validity of the study and concept mapping instrument was analysed. Emphasis was put on what was actually studied. Due to the questionnaire and focus question-based and concepts given concept maps, the analysis brought out, that the quality of concept maps does not depend on concept maps creating frequency and computer handling skills. The conclusion is that the concept mapping instrument was valid for this study.

To go further with this large-scale study, it would be interesting to compare the results with the scenario-based three-dimensional PISA-liked knowledge test and the results on the concept mapping part. It would give another unique possibility to interpret the results of this large-scale concept mapping research.

5. References

Cañas, A. J., Bunch, L. & Reiska, P., (2010) CmapAnalysis: an Extensible Concept Map Analysis Tool. In J.Sánchez, A. J. Cañas, J. D. Novak (Eds.), *Concept Maps: Making Learning Meaningful, Proc. of the Fourth Int. Conference on Concept Mapping*, Viña del Mar, Chile: Universidad de Chile.

Cañas, A, Novak, J. N, Reiska P (2012) Freedom Vs.Restriction of Content and Structure During Concept Mapping. Possibilities and Limitations or Construction and Assessment. A. J. Cañas, J. D. Novak, J. Vanhear (Eds.), *Concept Maps: Theory, Methodology, Technology, Proc. of the Fifth Int. Conference on Concept Mapping,* Valletta, Malta.

Chang, K.-E., Sung, Y.-T., Chang, R.-B., & Lin, S.-C. (2005). A New Assessment for Computer-based Concept Mapping. *Educational Technology & Society,* 8 (3): 138-148.

Gouli, E., Gogoulou, A., & Grigoriadou, M. (2003). A coherent and integrated framework using concept maps for various educational assessment functions. *Journal of Information Technology Education,* (2) 215-239.

Novak J. D. (2010). *Learning, Creating and Using Knowledge: Concept Maps as Facilitative Tools in Schools and Corporations.* New York: *Routledge* .

Reiska, P., Cañas A. J., Novak J. D., & Miller, N. L. (2008). *Concept mapping for meaningful learning and assessment The need for a paradigm shift in science education for post-soviet societies.* Frankfurt: Peter Lang Company (128-142).

Ruiz-Primo, M. A., & Shavelson, R. J. (1996). Problems and issues in the use of concept maps in scienceassessment. *Journal of Research in Science Teaching*, 33 (6), 569-600.

Ruiz- Primo, M. A, Schultz, S. E & Shavelson, R.J. (1997). *Concept map-based assessment in science: two exploratory studies.* University of California, Los Angeles: technical report.

Soika, K., Reiska, P., & Mikser, R. (2012). The Concept Mapping as an Assessment Tool in Science Education. A. J. Cañas, J. D. Novak, J. Vanhear (Eds.), *Concept Maps: Theory, Methodology, Technology, Proc. of the Fifth Int. Conference on Concept Mapping,* Valletta, Malta.