JOURNAL OF BALTIC SCIENCE EDUCATION

ISSN 1648-3898

USING CONCEPT MAPPING FOR ASSESSMENT IN SCIENCE EDUCATION

Katrin Soika,
Priit Reiska

Abstract. *Conceptual acquisition is an important aspect of science education, especially with regard to committing things to long term memory. Three case studies were carried out to examine approaches to conceptual acquisition in science education. The aim of these studies was to determine whether concept mapping method is a valid research method. To answer the question, studies were developed. Two of them used animation to describe abstract topics in chemistry. The aim of the third study was to compare the differences in higher taxonomy concept maps, created in an allocated timeframe.*
The findings of the studies indicate that the use of voiced interactive animation without teacher explanation does not lend itself to superior concept acquisition. As a result of the study it was concluded that concept mapping as an assessment method provides a unique possibility to visualize the structure of students' conceptual achievements. Based on the outcomes it is recommended concept mapping should be used more as an assessment or research method.
Key words: *assessment, concept mapping, animation, paper-based instruction, concepts.*

Katrin Soika, Priit Reiska
Tallinn University, Tallinn, Estonia

Introduction

Conceptual acquisition is an important component of science education, yet many such concepts are abstract and difficult to acquire. Determining the level of science learning, however, is handicapped if the assessment tool is too contingent on memorisation, especially short term memory. In this case, concept acquisition is poorly assessed since the assessment tool is more likely to measure memorisation powers.

The aim of the research was to seek evidence as to whether the use of concept mapping as an assessment tool can be helpful in measuring meaningful concept acquisition.

Concept Mapping

Concept mapping as an assessment method was developed by Joseph Novak and his research team in the early 1970s (Novak, 2010). The method is based on the theory of Ausubel (1968)-also called meaningful learning- and assumes that learners construct their knowledge, influenced by the knowledge which they have already acquired previously. This influence thus suggests that connections can be established between the prior learning and new learning and that these connections can be expressed by means of a graphical display, usually called a concept map. A concept map therefore links concepts, with prior concepts forming the base and the new conceptual acquisition building on this base, not necessarily in a one-to-one linkage system.

In a concept map, concepts- which could be words, things, pictures, formulas, symbols, etc- are linked with lines, labelled as proposition (Reiska, Cañas, Novak & Miller, 2008; Ruiz-Primo, Schultz & Shavelson, 1997; Chang, Sung, Chang & Lin, 2005). As such, the concept map is a collection of propositions, which explain the connections between the concepts. When constructed in a certain way, they can graphically express meaningful inter-relationships, existing between one or more different concepts (Ruiz-Primo et al., 1997).

Concept maps can be widely used in the school learning process. Research shows that the method helps to prevent rote learning, enables to summarize previously studied concepts or those evolving from class discussions (Novak, 2010), can be helpful in creating presentations, etc. However, concept maps can also be used as an assessment tool, being superior to other assessment tools in giving an overview of a student's mental processes and their structure (Gouli, Gogoulou & Grigoriadou, 2003).

Concept Mapping for Assessment

Novak pointed out that concept maps are perhaps best used as an assessment tool. Novak and Cañas (2006) wrote in their online technical report (p. 27): „There is nothing written in stone that says multiple choice tests must be used from grade school through university, and perhaps in time even national achievement exams will utilize concept mapping as a powerful assessment tool..." Science studies have proved the necessity of concept maps in assessing. For example: Walker and King (2003) concluded from two pilot studies investigating the use of concept mapping for assessing students' knowledge at a given point and over time that: a) concept mapping was a useful tool for building a portrait of the process of knowledge transformation from novice to expert; b) concept mapping was an appropriate tool for student assessment.

Some authors have pointed out weaknesses in the concept mapping method as an assessment tool. They argue that creating an acceptable structure of concept maps could be difficult for a novice. Furthermore, it could be challenging for instructors to evaluate the results (Chang et al., 2005), and the validity and reliability of concept maps have been questioned (Ruiz-Primo & Shavelson, 1996; Ruiz-Primo, 2004).

Klassen (2006) found in his research (p. 842): "Even though concept maps, both in instruction and in assessment, are able to demonstrate strong effects, it is not certain how to interpret the results. Also, because of a lack of standardization of scoring methods, reliability is difficult to achieve. In order for concept map assessment to gain legitimacy for large-scale usage, these issues must first be overcome."

Using and Interpreting Concept Maps

Neither creating nor evaluating the use of concept maps has been an easy task. Initially, concept maps were created mainly by pen and paper and results were calculated manually. The whole process was long and tedious. Nowadays, it has become much easier because of the use of computers (Novak, 2010). There are several measures for analysing concept maps ranging from the number and quality of propositions, to the size and hierarchy of the concept map, or clusters of maps. To evaluate concept maps, specific dimensions for measuring are needed. Miller, Cañas *et al.* (2006) have developed a topological taxonomy for evaluating concept maps. They considered five criteria, through which topological levels were defined: 1) recognition and use of concepts; 2) presence of linking phrases; 3) degree of ramification; 4) hierarchical depth, and 5) presence of cross-links (Reiska, Cañas, Novak, & Miller, 2008). The taxonomy consists of 7 levels: identified as from 0 to 6. The maps which are classified as 5 and 6 meet all or almost all of the criteria.

Trumpower and Sarwar (2010) acclaimed that concept maps hold promise for being used effectively in formative assessment systems. Since formative assessment is used as an indicator of progress, it must identify a student's strengths and weaknesses, while also being easy to use. Fortunately, computers have made it easy to create and assess such concept maps, allowing teachers and researchers to use this method in their everyday work.

The concept mapping method was also used as an assessing method in an extensive longitudinal study Lotegüm, conducted in Estonia, where 1614 students were examined with exercises that involved concept mapping and PISA like three dimensional scenario-based exercises. Currently, the evaluation process is ongoing as experts are marking the propositions. The preliminary results of students' concept maps derived from general data and answers to the question list are collected (the research examined the frequency of creating concept maps and problems occurring during the study). It turned out that scores of the concept maps (taxonomy score and proposition count) depend neither on the frequency of creating concept maps nor computer handling skills (Soika, Reiska, 2013). The preliminary results were analysed by *CmapAnalysis* software which was created by Cañas, Bunch and Reiska (2010).

CmapAnalysis Program

This program allows analysis of various algorithms, rubrics and techniques used within concept maps. Parameters are defined by the researcher. The creators of the program propose that this software helps instructors, researchers and teachers to automate routine analytical operations (Cañas, Bunch & Reiska, 2010).

The CmapAnalysis software supports: a) taking input Cmaps into the open CXL file format in addition to the Cmap format, thus allowing the analysis of concept maps developed by concept mapping programs that utilize CXL, b) enabling users to add other measures to the program.

CmapAnalysis gives an opportunity to measure different categories, for example:

a) size: describes how many concepts, linking words and propositions are included in a concept map;
b) quality: describes what kind of concepts, linking words and propositions are included in a concept map,
c) structure of concept maps: describes how well the concepts are inter-connected; determines clusters, identifies three most central concepts in each map, etc.

The usefulness of the program lies in the fact that it can handle large sets of concept maps and also the scores, which lend themselves to further analysis, for example using MSExcel (Cañas, Bunch & Reiska, 2010).

In this study, the following parameters from the concept maps were calculated:

a) proposition count: the number of propositions (i.e. concept-linking phrase-concept) included in the map,
b) proposition quality: expert evaluation of the appropriateness of a concept link and a numerical score or „mark" given to the proposition,
c) taxonomy score: calculated by different measures, including average words per concept, branch point count, concept count, linking phrase count, separated concepts count, proposition count, count of concepts with outgoing connections but no incoming connections (root child), sub map count. The taxonomy score is expressed with a number between 0 and 6, with higher scores typically indicating higher quality concept maps.

Animation

Animations, moving illustrated materials, are being used increasingly frequently in schools to depict dynamic changes over time and space, and illustrate phenomena or concepts that might be otherwise difficult to visualise (Nakhleh, 1992; Mayer & Moreno, 2002; Ruiz, Cook, & Levinson, 2009; Kozma & Russell, 2005). Studies from many countries have demonstrated positive effects in the use of different and innovative methodologies and visualisation technologies on students' understanding of central scientific concepts (Wu & Shah, 2004; Kozma & Russel, 2005; Nakhleh, 1992, Soika, 2007). Nevertheless, the effectiveness of the animation depends on the student's personality- it involves memory, the cognitive load (Mayer & Moreno, 2002), the structure of the animation used and the methodology of using the animation in the classroom (Ruiz et al., 2009; Wu & Shah, 2004; Mayer 2005).

Methodology of Research

The methodology, data collection, analyses and results of the case studies that used concept maps as an assessment tool will be explained individually in detail. The three studies were carried out in years 2010 to 2011. These studies had different research questions, but shared concept mapping as the method for collecting evidence. A previous study (Soika, Reiska & Mikser, 2010) pointed out that the concept mapping method allows a much better analysis of the structure of students' knowledge than questionnaires. The data collection was structured as indicated in Figure 1 below.

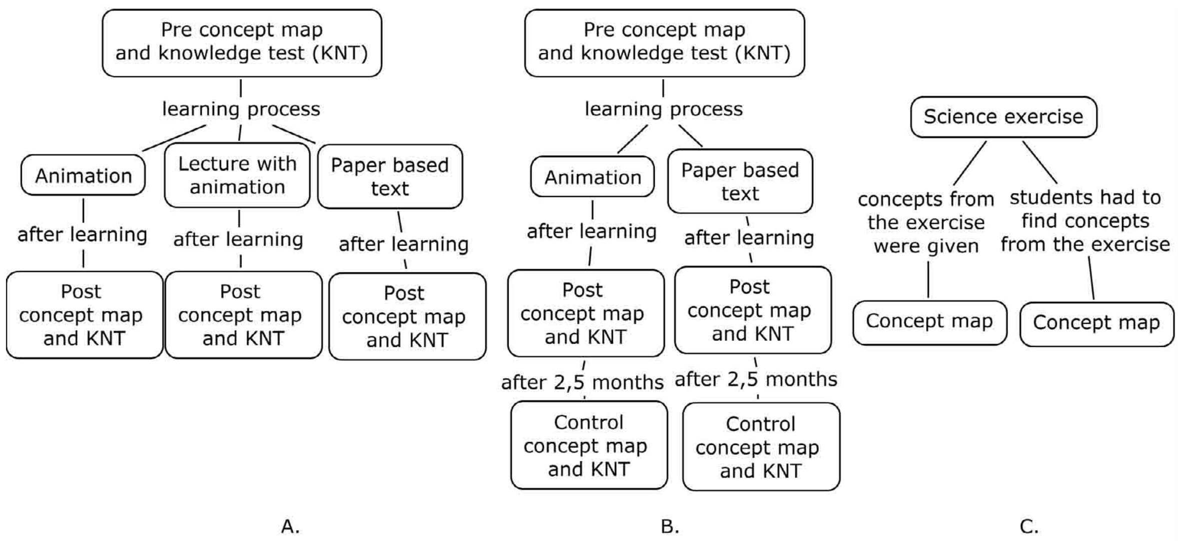

Figure 1: Structure of data collection in three different case studies.

The animations used had a voice-over explanation; the paper-based text had identical information to the animation conveyed in words and pictures.

The three studies had different aims. Studies A and B analysed how concept maps could be used as an assessment tool for evaluating students' knowledge. Both studies also featured ordinary open-ended questions and multiply choice questions which allowed comparing the results. The most significant differing parameter of the two studies was the time factor.

First Case Study- Using Concept Mapping for Analysing Different Learning Methods
Case 1

The general aim of this study was to investigate the kind of impact that animation has on students' knowledge. The research questions for this case were:

1. What kind of impact does animation have on the knowledge of students studying individually?
2. Are there any differences in students' knowledge if the same animation is explained by a teacher?
3. Do different groups of students create similar concept maps?
4. Which students create higher valued propositions: those who study through animation, or those who read paper-based texts?

These questions were based on the outcomes of a previous study (Soika, Reiska & Mikser 2010), where concept maps analysis led to the conclusion that students from various groups had created dramatically different concept maps, which seemed to depend on the studying method. It was predicted that students learning from animation would create higher scored maps and knowledge tests.

Data Collecting and Analysing

In this study, 77 students were divided into three groups (the first group studied individually using animation with voice explanation; the second group had the animation supported by teacher explanation; the third group studied the same topic individually from a paper-based text). The topic of the study materials was identical and new for students. The main data collecting method is summarized in Figure 1. Scheme A.

Students were asked to create two concept maps using 20 given concepts in an allocated timeframe. Both pre- and post-concept maps were analysed using the *CmapAnalysis software*, although the content and correctness of the sentences were assessed manually. Students also gave answers to pre- and post-learning questionnaires (knowledge tests). The results were compared in MSExcel.

ISSN 1648–3898

Second Study- Using Concept Mapping for Assessing Long-Term Memory
Case 2

Derived from the outcome of the first study, a new study was planned. The research questions for the second case were:

1. What kind of impact does a longer time period have on knowledge-based concept maps?
2. Are there any differences in knowledge of various groups of students?
3. What kind of similarities can be found in concept maps of various groups?

Data Collecting and Analysing

In this study 62 students were divided into two groups- one group studied individually from animation with voice explanation and the other studied the same topic individually from a paper-based text. The study materials were identical, including the same text and pictures. The main data collection method is shown in Figure 1. Scheme B.

Students were asked to create three concept maps in the allocated time from 22 concepts, which were given to them. Pre-concept maps (made before learning), post-concept maps (made on the same day after learning) and control concept maps (created 2,5 months after learning), were analysed with CmapAnalysis software, the content and correctness of the sentences were assessed manually. Students also answered to pre-, post- and control- questionnaires (knowledge tests, with seven questions) and a "pleasantness" inquiry. The results were compared in MSExcel.

Third Study- Assessment with Concept Mapping in Different Conditions
Case 3

The aim of the third study was to compare differences in concept maps created by students who have pre-given concepts to work with and students whose task it is to find concepts on their own from the study material. This time data was collected from students with the help of an exercise of natural science.

The research questions for the third case were:

1. What kind of concepts do students identify from a narrative science exercise? Are these concepts similar to concepts identified for the same exercise by a teacher?
2. Do different conditions lead to similar learning outcomes? The conditions were:
   a) students had to create a concept map in the allocated time from 20 pre-given exercise- based concepts of various level of difficulty,
   b) students had to find 20 concepts from the exercise and to create a concept map in the allocated timeframe.

Data Collection and Analysis

For the third study, 54 students were divided into two groups. Both groups received the same science exercise. Students had to read the text and find answers to multiple questions on various difficulty levels. Having done that, they created a concept map, which had to describe natural science and everyday life, connected to the meaning of the exercise (Figure 1. Scheme C).

The exercise described an industrial accident which happened on 4 October 2010 in Hungary, when toxic "red mud" spill was released, eventually reaching the Danube, killing at least 9 and injuring 150 people. There were 4 science-based questions in the exercise representing different levels of difficulty.

Students had 20 minutes to read and solve the exercise and another 20 minutes to create a concept map about the exercise. Maps were analysed manually as well as using CmapAnalysis and MSExcel software.

Results of Research

Case 1

To give a better overview from the results of the first case, Figure 2 depicts the changes in a diagram.

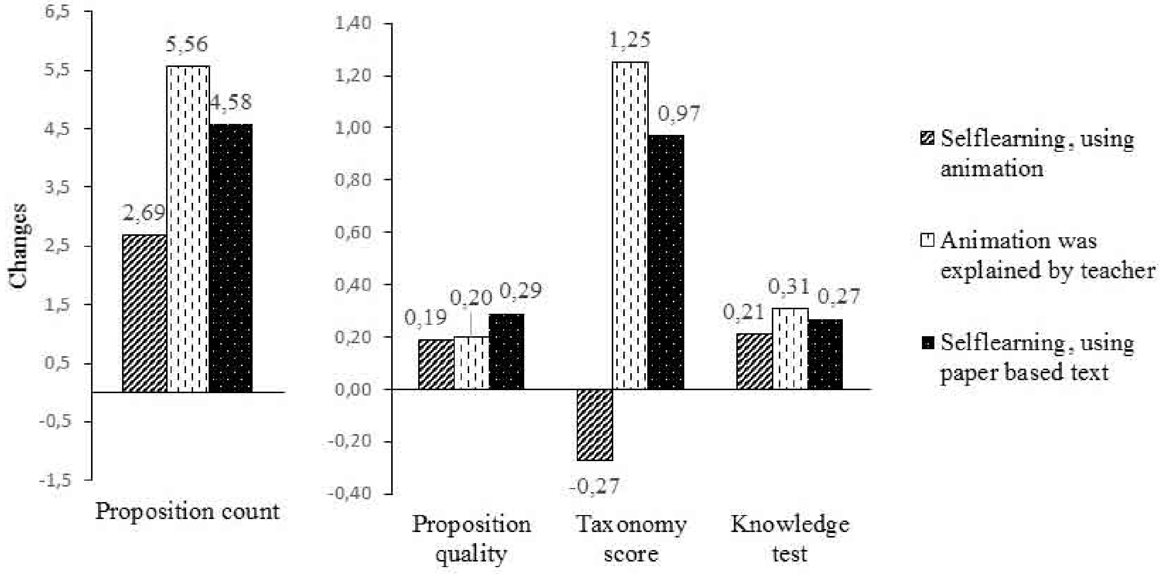

Figure 2: Changes in different measures (first case study).

Results show that the highest quality concept maps were created by the group who studied using animation with teacher explanations, although these explanations were identical to the voice explanation which appeared in the animation. The group who studied from the animation with a teacher received similar results to the group who studied individually from the paper based text. The group who studied individually from the animation had different results.

The results of this study point out that the structure of created concept maps depends on the studying method- although such difference is not indicated in a knowledge test.

1. The results of students learning individually from paper are better than the results of students learning individually from animation. The average taxonomy score of the group who studied individually from the animation was actually negative, since the scores dropped- students created more poorly structured concept maps after the learning process than they had before.
2. On average, the concept maps of students learning individually from the animation had the lowest number of propositions created after the learning process. Concept maps created before and after the learning process did not differ much in terms of number of sentences.

ISSN 1648–3898

Case 2

The results were analysed in *CMapAnalysis* and *MSExcel*. The most important results of the study are shown in Figure 3.

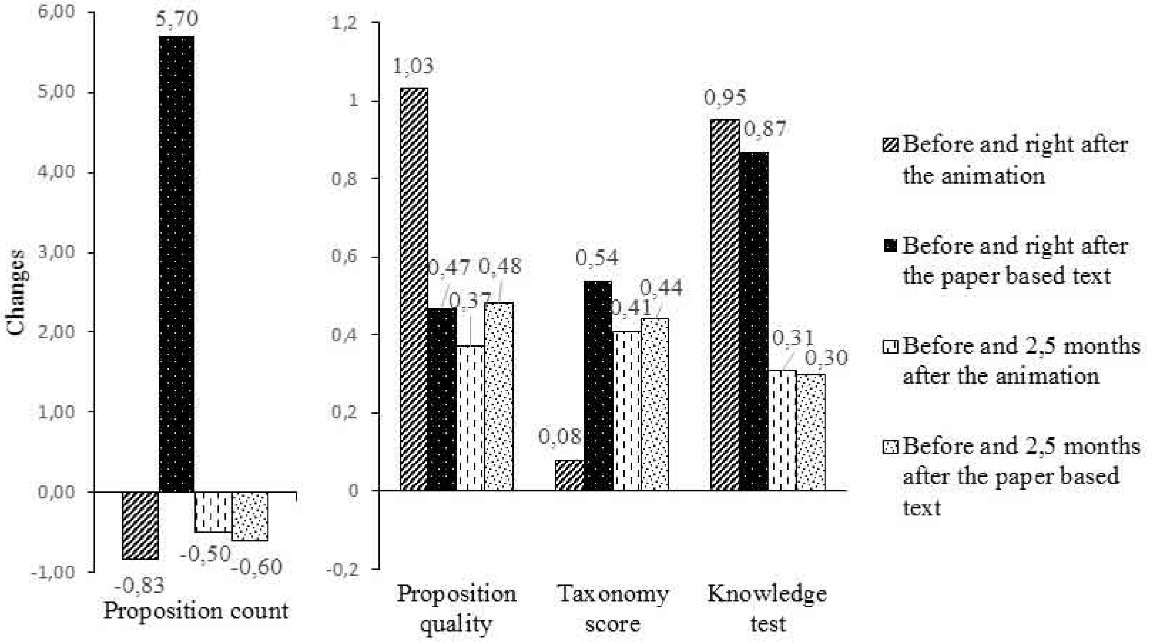

Figure 3: Average changes in measures per students (second case study).

After learning from the animation, concept maps of students were similar to each other and did not include as many propositions as the maps made after learning from the paper. It was obvious that studying method had impacted connections between concepts. Question-based knowledge test was conducted before and after the learning process, but discovered no significant differences of changes between groups. After 2,5 months control maps of both groups were similar again (Figure 3). There had been an influence in students' knowledge right after the learning process. Based on previous experience, the students, who had studied from the animation, were expected to remember more details about the abstract topic and would create extremely high-quality concept maps and give good answers to knowledge tests. The results, however, did not support this premise. A question arose: why did the students' concept maps not differ after animation-based learning? How is it possible that students who have learnt from paper-based instruction create topologically superior concept maps to those who have seen the animation? Students were asked to answer open- ended questions, where they described their emotions and advantages, connected with animations and paper-based instructions.

The aim of the open-ended questions, asked from students right after the learning process as well as 2,5 months later, was to find out how they liked the different types of instruction. All the students admitted that they had liked the animation whereas only a few claimed to have enjoyed the paper-based instruction. The students who had seen the animation 2,5 months earlier wrote that there were colours, they enjoyed the animation and movements- but they couldn't explain the idea of the abstract topic (a sample answer from the open-ended questionnaire: "*I remember colours and moving particles, but actually I do not remember why they were hanging around*…"). Those who had studied from the paper-based instruction, recalled reading the text with some pictures in it, but could not point out any main idea representing the topic, either.

Case 3

The aim of the study was to examine the differences in scores of higher taxonomy concept maps created in an allocated timeframe, by the students who find concepts by themselves from data provided in the exercise, compared with the students who are given relevant concepts by an instructor.

Conclusion of the analysis points out that the students who had previously received concepts to work with, created maps with higher taxonomy scores and more propositions than the other group: average taxonomy scores, respectively, were 3,75/ 3,03; average number of propositions per student 18,5 /16,6; in addition, visual differences emerge in the maps (Figure 5 and Figure 6). It occurred, that 14 of 16 most frequently represented self-determined concepts were also given by the instructor to the other group. So it was realised that in a limited timeframe, students create concept maps with higher taxonomy scores if concepts are given to them.

Calculating differences in the connections to three most central concepts provide us with the opportunity to analyse the structure of the entire map. As observing, the concept maps created by students who were not given concepts beforehand tend to take the shape of a "star" whereby a single concept lies in the middle while others are connected to it mainly by one label. When 20 concepts were given to students, connections between concepts created more of a network (illustrative diagram in Figure 4, illustrative concept maps in Figure 5 and Figure 6).

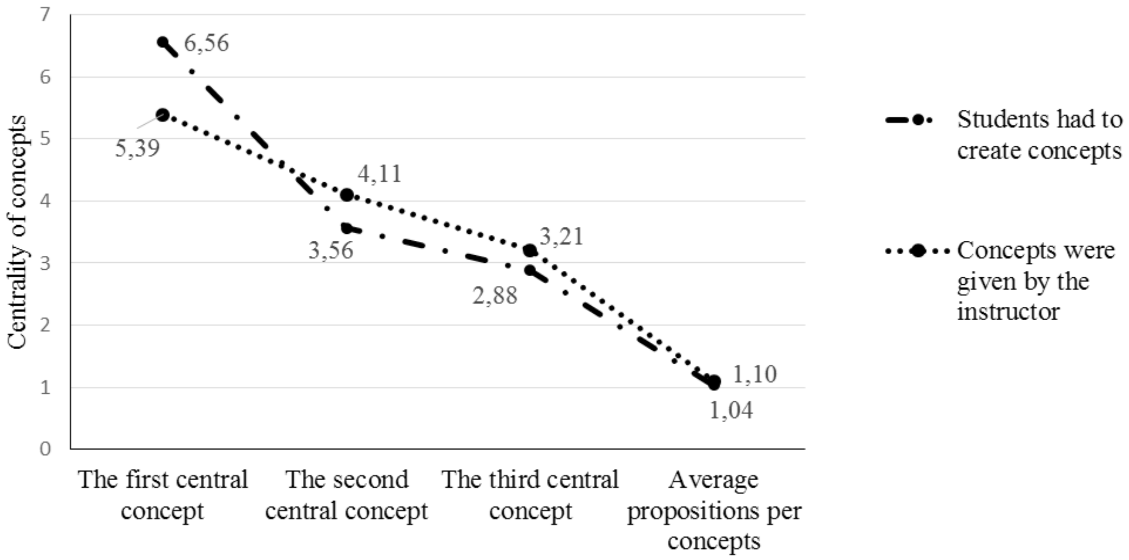

Figure 4: Differences in number of connections to three most central concepts presented by the two groups of students.

When students had to find concepts on their own, the centrality of the first concept is especially dominant, producing high average of propositions per concept in a "star"-like formation (Figure 5). However, since the concepts do not generate integrated entirety, the taxonomy score remains low. The concept maps with more inter-linked concepts provide higher taxonomy scores, but fewer connections with the most central concept (Figure 6).

ISSN 1648–3898

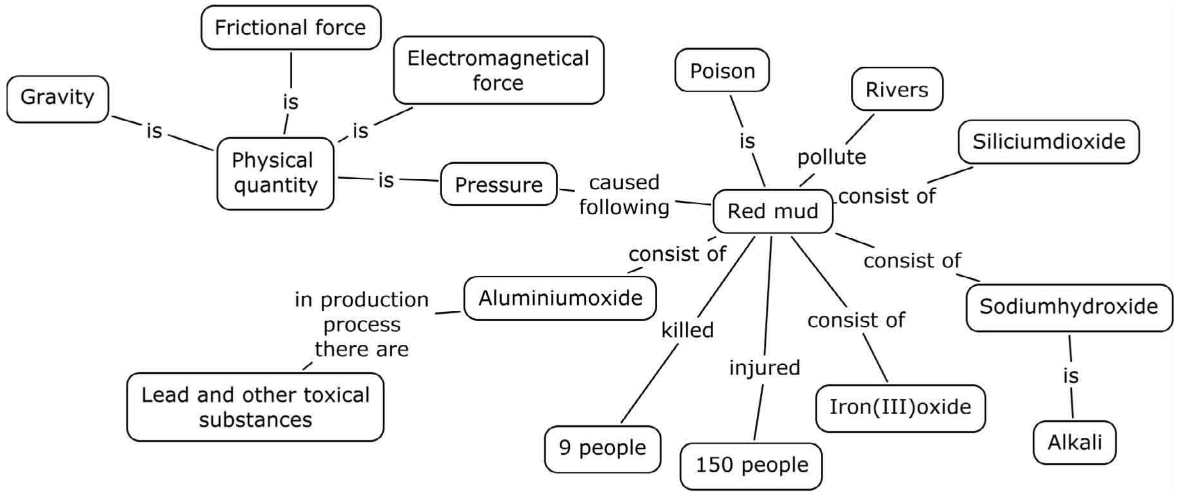

Figure 5: **A concept map, constructed by a student who had solved the science exercise and was asked to find concepts by herself from the exercise.**

The first central concept is "red mud" and the second central concept is "physical quantity", with the first central concept being the clear dominant. There are 15 propositions and 16 concepts in the concept map and the taxonomy score of the map was 2.

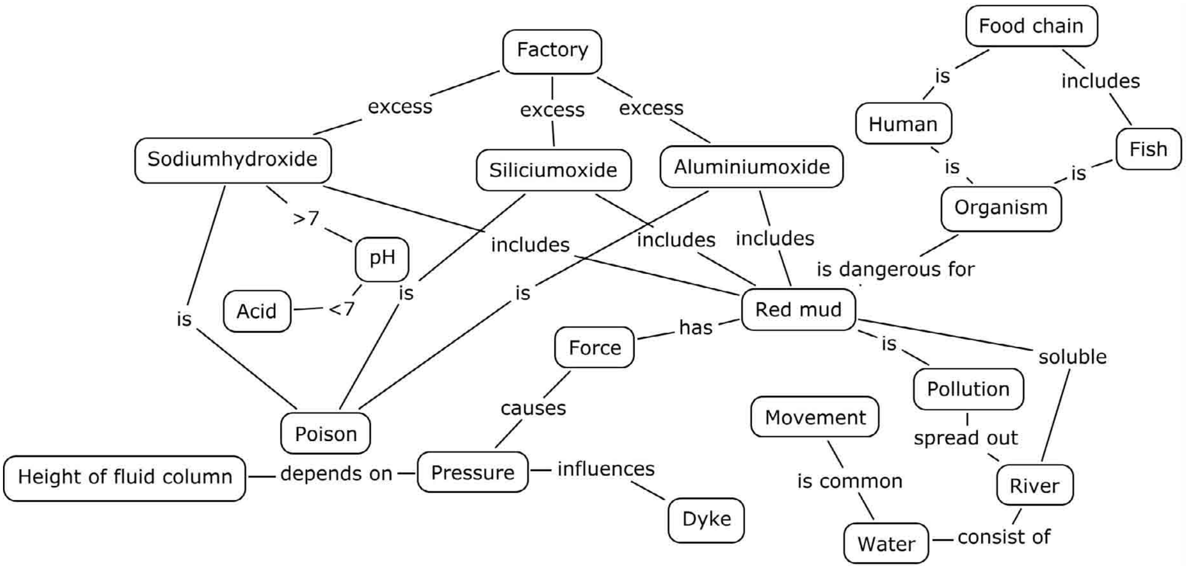

Figure 6: **A concept map, constructed by a student who had solved the science exercise and received concepts about the exercise from the instructor.**

All of the concepts presented in Figure 6 were given to students. There are 20 concepts and 25 propositions in the concept map and the taxonomy score of the map was 4. Figure 5 and Figure 6 have different structures. As observing Figure 6, this is characteristic of the maps generated by the second group of students.

Discussion

The first study manifested: 1) it is good to evaluate changes in students' knowledge with concept maps because ordinary knowledge tests do not reveal a similar process in students' knowledge, 2) abstract animation should be used in lectures if supported by teacher's explanations. Based on previous experiences it could be said that students do not properly focus on content if they are watching animation individually from a video. Mayer (2005) and other researchers have emphasized that while learning from multimedia instruments, should be remembered that cognitive channels must not be overloaded. Students should not receive too much different information (voice, colours, flashing screen, movable particles etc) from the study material, because learning is a very complicated process. The students who followed the animation with teacher explanations were guided and did not forget the aim of their work.

The outcomes of the second case, that was derived from the outcome of the first case, indicate that abstract animation had no significant effect on students' long-term memory- the results of tests were almost equivalent. Following the law of forgetting of Ebbinghaus that shows how people forget what they memorize, García and his research team (2007, p. 622- 623) wrote: "The law marks that irremediably makes us forget content, also holds a pleasant surprise: when the content is revised and a command of the subject is achieved for a second time, some parts that have already been learnt will be lost again, but the gradient of the forgetting curve it is not as high this time as it was after the first study". It is argued that students, who learnt the topic for the first time and had not dealt with it for the past 2,5 months, essentially did not remember the topic. According to this theory, students would have had to recall the previously studied topics in order to understand and remember. The abstract chemistry topic represented in current study was not reiterated in class for 2,5 months after it was studied first. Figure 3 pointed out, the changes in students' knowledge before learning and after the 2,5-month period are almost equal in concept maps and knowledge tests. It could be concluded that neither of the study methods had a competitive advantage over the other in committing things to students' long-term memory. The results were different right after the learning process. The students who had studied from the animation, created propositions (sentences) with higher scores, but made fewer propositions than those who studied from the paper-based text. Based on previous study (Soika et al 2010), the students, who had learnt from the animation, expected to score higher for the instruction-based propositions. Both the animation and the paper-based text presented a certain scheme from abstract chemistry concepts. The previous study (Soika, 2010) demonstrated that the students, who had learnt from the animation, repeated almost the same animation-based scheme in their concept maps and added a few sentences on their own. By doing so, they received higher marks for abstract chemistry-based sentences, but created fewer propositions than before the studying process. The same dilemma repeated itself in the second case study of this research. However, the impact of the effect disappeared within 2,5 months.

Without the concept mapping method such a drastic change in students' mind could not have been noticed. Connections between the concepts visualised the fact that something had impacted the students, whereas after time the impact had disappeared.

After the analysis of the second study a question arose: what kind of concept maps would the students have created if they had had to find concepts on their own from the study material? Hence, the third study was conducted, following the second one.

The third study revealed that while using concept mapping to assess students, the aim of the work should be kept in mind, since this is what the result depends on. Teachers and researchers should always observe the need to use valid and reliable methods in assessing their students. The current study demonstrated that scores of concept maps also depend on the established conditions. A student, who read the exercise and found concepts from the instruction would have created a different concept map, had the concepts been given to him beforehand. Based on previous experience, it could be said that time is a limiting factor in generating concept maps. Students create more highly structured concept maps if they have time to think. Searching for concepts requires time and effort, so after the concepts have been found, students need additional time to create a high-quality concept map. Additionally, the third study pointed out that instructor-given concepts saved time and effort and gave students the opportunity to create concept maps that would generate higher scores.

Among the weaknesses of the concept mapping method that some authors (Ruiz-Primo & Shavelson, 1996; Ruiz-Primo, 2004; Chang et al., 2005) have pointed out, one is the assertion that creating an acceptable structure of concept maps can be difficult for a novice. Nevertheless, the method was used in case studies of the research. Also, recent study (Soika & Reiska, 2013) pointed out that neither computer handling skills nor the frequency of

ISSN 1648–3898

creating concept maps affect the structure and quality of concept maps created. These outcomes lead us to the conclusion that concept mapping is a valid method which should be used more regularly in research and assessment activities.

Klassen (2006) pointed out that when concept mapping method is used for assessment, the interpretation of the results is very important. Case studies of this research proved that in addition to the interpretation of the results, it is also very important to consider, how concept mapping method is used for assessment, e.g. are the concepts pre-given or not.

The three studies presented in this research, demonstrated that concept mapping method can be used as an assessment tool, since it details the structure of students' knowledge in a way that cannot be reproduced with other instruments, such as questionnaires.

Conclusions

In 2011, a new national curriculum for basic and grammar schools was introduced in Estonia (Gümnaasiumi riiklik õppekava, 2010). According to the curriculum, new aims and competences should be developed at school. The curriculum also requires teachers to use new methods, activities and information technology tools in the classroom. First two cases indicated that in order to improve the understanding and memorising of abstract processes, which are visualised in animations, valid methodology should be used by teachers.

New methods are needed for developing students' knowledge and directing them to create connections between the previously learned and new facts and principles. These competences should not be assessed with ordinary knowledge tests. Three cases revealed that concept mapping method is valid and reliable, provided the instructors acknowledge the aim of their work, and as such could be used more extensively at schools in learning and assessing processes. Scores of concept maps depend on several factors and instructors should factor them in while planning their work. Concept mapping method as an assessment tool is indispensable and gives a unique possibility to visualise the structure of students' knowledge even in assessing large scale studies.

Acknowledgements

This study has been supported by European Social Fund programme Eduko grant Lotegüm.

References

Cañas, A. J., Bunch, L., & Reiska, P. (2010). CmapAnalysis: An extensible concept map analysis tool. In. *Concept Maps: Making Learning Meaningful, Proceedings of the Fourth International Conference on Concept Mapping* (pp. 75-82). Viña del Mar, Chile: Universidad de Chile.

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.

García, R. R., Quirós, J. S., Santos, R. G., González, S. M., & Fernanz, S. M. (2007). Interactive multimedia animation with macromedia flash in descriptive geometry teaching. *Computers & Education*, *49* (3), 615-639.

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* (1), 215-239.

Gümnaasiumi riiklik õppekava (2010). Vabariigi Valitsuse 28. jaanuari 2010. a määrus nr. 13. [National curriculum for upper secondary schools Regulation No. 13 of the Government of the Republic of 28 January 2010]. Retrieved from: http://www.ibe.unesco.org/curricula/estonia/ er_usfw_ 2011_eng.pdf

Klassen, S. (2006). Contextual assessment in science education: Background, issues, and policy. *Science Education*, *90* (5), 820-851.

Kozma, R., Russel, J. (2005). Multimedia Learning of Chemistry *Cambridge handbook of multimedia learning.* (pp. 409-428). New York: Cambridge University Press.

Mayer, R. E., & Moreno, R. (2002). Animation as an aid to multimedia learning. *Educational Psychology Review, 14* (1), 87-99.

Mayer, R. E. (2005). Cognitive theory of multimedia learning. *The Cambridge handbook of multimedia learning*. (pp. 31-48). New York: Cambridge University Press.

Nakhleh, M. B. (1992). Why some students don't learn chemistry: Chemical misconceptions. *Journal of Chemical Education, 69* (3), 191-196.

Novak, J. D., & Cañas, A. J. (2008). The theory underlying concept maps and how to construct and use them. *Florida Institute for Human and Machine Cognition Pensacola Fl. Retrieved from: http://cmap.ihmc.us/Publications/ResearchPapers/TheoryCmaps/TheoryUnderlyingConceptMaps*.

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.* (pp. 128-142). Frankfurt: Peter Lang Company.

Ruiz J. G., Cook D. A., Levinson A. J. (2009). Computer animations in medical education: a critical literature review. *Medical Education*, *43* (9), 838–846.

Ruiz-Primo, M. A., & Shavelson, R. J. (1996). Problems and issues in the use of concept maps in science assessment. *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*. Center for Research on Evaluation, Standards, and Student Testing, Graduate School of Education & Information Studies, University of California, Los Angeles. Retrieved from: http://www.cse.ucla.edu/products/reports/TECH436.pdf

Ruiz-Primo, M. A. (2004). Examining concept maps as an assessment tool. In *Concept maps: Theory, methodology, technology. Proceedings of the First International Conference on Concept Mapping*. Pampolonia, Spain: Universidad Publica de Navarra. Retrieved from: http://cmc.ihmc.us/papers/cmc2004-036.pdf

Soika, K. (2007). Infotehnoloogia võimaluste kasutamine metoodiliste vahendite väljatöötamiseks gümnaasiumi orgaanilise keemia reaktsioonimehhanismide näitlikustamiseks ja ainete nimetamise algoritmi koostamiseks. [Using information technology for illustrating the reaction mechanisms in high school organic chemistry classes and algorithm for nominating substances.] Tallinn: Tallinn University.

Soika, K., Reiska, P., & Mikser, R. (2010). The Use of Animation as a Visual Method in Learning Chemistry. In *Concept Maps: Making Learning Meaningful, Proceedings of the Fourth International Conference on Concept Mapping*. Viña del Mar, Chile: Universidad de Chile. Retrieved from: http://cmc.ihmc.us/cmc2010papers/cmc2010-b12.pdf

Soika, K., Reiska, P., & Mikser, R. (2012). The Concept Mapping as an Assessment Tool in Science Education. *Proc. Of the Fifth Int. Conference on Concept Mapping*. Valletta, Malta: University of Malta. Retrieved from: http://cmc.ihmc.us/cmc2012papers/cmc2012-p188.pdf

Soika, K., Reiska, P. (2013). Large Scale Studies with Concept Mapping. *Journal for Educators, Teachers and Trainer, 4* (1), 142 – 153.

Trumpower. D. & Sarwar, G., S. (2010). Formative Structural Assessment Using Concept Maps as Assessment for Learning. In *Concept Maps: Making Learning Meaningful, Proceedings of the Fourth International Conference on Concept Mapping.* (pp. 132- 136) Viña del Mar, Chile: Universidad de Chile.

Walker, J. M. T., & King. P., H. (2003). Concept mapping as a form of student assessment and instruction in the domain of bioengineering. *Journal of Engineering Education*, *19* (2), 167-179.

Wu, H. K., & Shah, P. (2004). Exploring visuospatial thinking in chemistry learning. *Science Education**, 88*** (3), 465 – 492.

Received: *April 29, 2014*

Accepted: *August 20, 2014*

Katrin Soika	PhD Student (MSc), Tallinn University, Narva mnt 25, 10120 Tallinn, Estonia. E-mail: katrins@tlu.ee
Priit Reiska	Dr. paed. habil., Professor for Science Education, Vice Rector, Tallinn University, Narva mnt 25, 10120 Tallinn, Estonia. E-mail: priit.reiska@tlu.ee