UNIVERSITEIT TWENTE.

Final Project Thesis

Developing a Tool for Learning Concept Maps

Author:
M.C. van den Enk
[s1004654]
m.c.vandenenk@student.utwente.nl

Supervisor:
dr. A.H. Gijlers
a.h.gijlers@utwente.nl

June 29, 2016

Contents

Part I

Design

Part II

Research Proposal

Summary

In literature, a distinction is often made between meaningful encoding and rote learning in terms of retrieval practices, of which the former is regarded to be more beneficial than the latter. A prominent example of meaningful encoding is the concept map, where retrieval practice is often used by means of flashcards. However, recently studies have indicated the importance of both aspects. Because of this, a new tool is developed, attempting to bridge the gap between meaningless rote memorisation and meaningful learning by combining the active retrieval mechanisms of flashcards with the meaningful visualisation by concept maps. In the proposed research, the effects of flashmaps will be investigated in terms of the learning effects measured by learning gain, and the affective effects measured by a questionnaire. Furthermore, the use of the tool will be investigated by means of interviews and user logs. The research will take place within a Dutch classroom setting by first measuring prior knowledge with a pre-test, then letting the students use the tool for 7 days and finally measuring their final knowledge with a post-test. Finally, a sample of the students will be interviewed. This research hopefully leads to a better way of meaningful rote memorisation and gains more insight in how flashcard systems and concept maps can benefit from each other.

Project Description

Problem Statement

Both the taxonomy of learning by Bloom, Engelhart, Furst, Hill, and Krathwohl (1956) as a revision of this taxonomy by Krathwohl (2002), as well as the three stages of skill acquisition by Anderson (1982), propose that all learning should start with memorising factual knowledge. Furthermore, von Glaserfield (2001), one of the main founders for critical constructivism, expresses a need for training students so that they permanently possess facts and are able to repeat them flawlessly whenever they are needed, while also understanding what is placed into their memory. Ericsson and Kintsch (1995) adds to this by stating that in order to perform complex tasks, people must maintain access to large amounts of information, and that solely encoding knowledge is not sufficient. Despite this, Karpicke (2012) argues that "[r]etrieval processes, the processes involved in using available cues to actively reconstruct knowledge, have received less attention" (p. 158), whereas basic research on learning and memory has emphasised that retrieval must be considered in any analysis of learning. Therefore, this project aims to research a tool for meaningfully enhancing the retrieval process.

Karpicke (2012) also states that meaningful learning often is defined in contrast to rote learning, and that active retrieval is thought of as an example of the latter leading to poorly organised knowledge that lacks coherence and integration. However, in another study they found active retrieval to enhance learning of meaningful educational materials and that these effects are long-lasting, not short-lived (Karpicke & Blunt, 2011). In this study, he compared the effects of active retrieval using measures of meaningful learning contrasting to a popular learning strategy known as concept mapping. The latter involves a graph consisting of nodes representing concepts and labeled lines denoting the relation between a pair of nodes (Ruiz-Primo & Shavelson, 1996) (see figure 1). Multiple researchers have found by means of both qualitative and quantitative studies that concept maps can promote meaningful learning leading to positive effects on students (Hwang, Wu, & Ke, 2011; Subramaniam & Esprivalo Harrell, 2015; Canas & Novak, 2012). This has been demonstrated in comparison to activities such as reading text passages, attending lectures, and participating in class discussions (Singh & Moono, 2015; Nesbit & Adesope, 2006). Canas and Novak (2012) describes the process of concept mapping as the only effective way of using the concept map, which refers to students constructing their own concept maps. This is why the concept map is generally viewed as a tool in alignment with the constructivist perspective. Because of this, the concept map might seem as a solution to the need asked by von Glaserfield (2001) and his peers. However, the aforementioned article by Karpicke and Blunt (2011) reveals that retrieval practices produced better performance than elaborative concept mapping for meaningful learning.

One of the currently existing methods for efficiently rote memorising information is the flashcard system, which entails studying declarative knowledge using active retrieval in a so-called paired-associate format. Within this format, learners are asked to associate terms with other

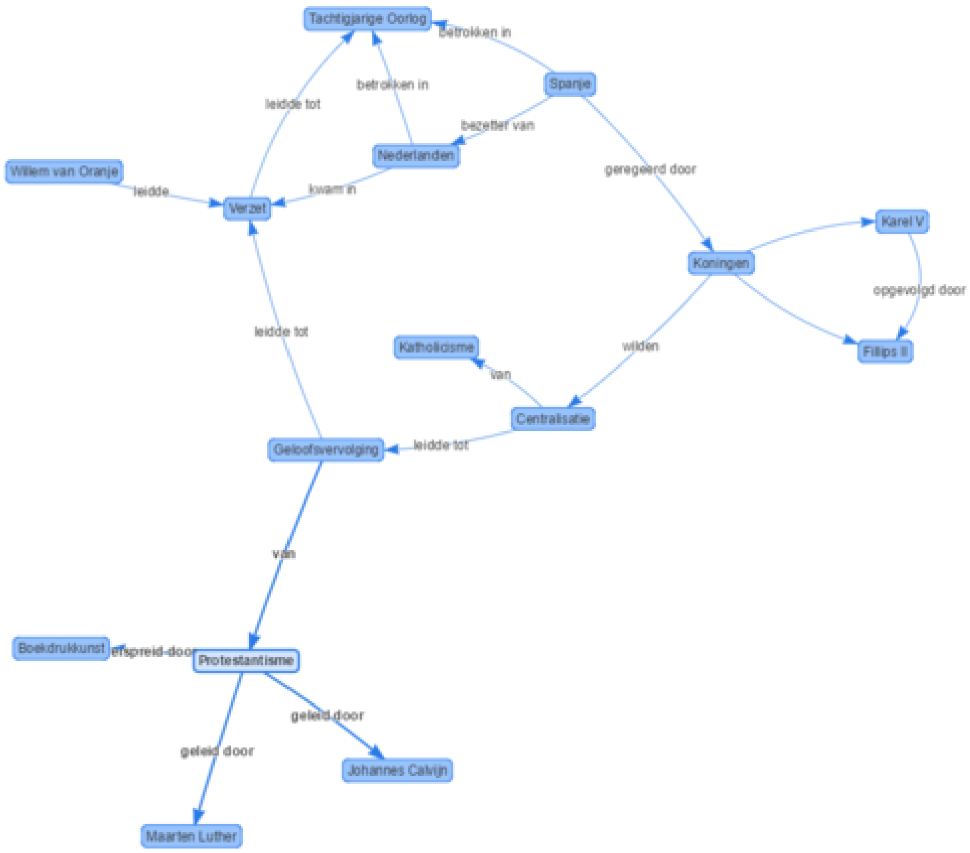

Figure 1: An example of a concept map

terms outside meaning-focused tasks (Nakata, 2011), for example by associating a definition with a presented concept. With flashcards, large numbers of words can be memorised in a very short time, and are more resistant to decay (Nakata, 2011; Joseph, Eveleigh, Konrad, Neef, & Volpe, 2012). Furthermore, when evaluating flashcards in a psychology setting, it was found that students who use flashcards have a significantly higher final average than those who do not (Burgess & Murray, 2014; Golding, Wasarhaley, & Fletcher, 2012).

Per contra, not all research favours using flashcards for textual comprehension. Zirkle and Ellis (2010) and iMcCullough (1955) state that flashcards are especially useful for learning declarative knowledge but not for textual comprehension. Zirkle and Ellis (2010) points out the overemphasis placed upon the rote memorisation of disconnected facts, whereas whatever it is that students are to place into memory they should, more importantly, understand. Furthermore, Hulstijn (2001) describes flashcards as a relic of the old-fashioned behaviourist learning model, and states that we have to look for more modern constructivist models.

Solving these problems could lead to better utilisation by teachers and students of producing a store of knowledge that remains flexibly retrievable, in contrast to only segregated paired associations which depend on specific cues in order to be retrieved. Furthermore, using computer-based flashcards have been used very widely (Nakata, 2011; Burgess & Murray, 2014; Golding et al., 2012; Kornell & Bjork, 2008), and improving currently existing flashcards could reach a wide audience of future users of flashcard systems.

Therefore, another solution might be the development of a new tool, which will from henceforth be referred to as the flashmap system. The intention behind the flashmap system is to combine the paired associate mechanism of the flashcard system with the visual representation of the concept map, and is a new tool designed and developed for this research project. This tool might have the potential to bridge the gap between the two systems and therefore make meaningful and effective rote memorisation possible, for it makes the relations between the concepts explicit to the student and thereby increasing the organisation of the knowledge and reducing the segregation of facts. Thereby, it might provide a solution for the problems by Zirkle and Ellis (2010) described before.

For evaluating this flashmap system, a group of Dutch high school teachers of the Stedelijk Lyceum has been found willing to participate, with their students using either the flashmap or the flashcard system for self study parallel with classroom instruction. The content of the instruction will be the history of Dutch literature during the sixteenth and seventeenth century. For example, the students have to learn what the influence is of the Dutch War of Independence on the *Spaanschen Brabander* by Bredero. Because of the content existing mainly of concepts with meaningful relations it fits to the concept map technique and thereby the flashmap system could be significantly beneficial over the flashcard system.

In conclusion, flashcards systems are an effective tool for meaningful learning, but could be enhanced by visualising it with concept maps, and therefore the effects of using a flashmap system over using a flashcard system will be investigated.

Theoretical Conceptual Framework

Flashcards systems

The most straightforward example of a flashcard system is a deck of physical cards, with on one side a question and on the other side an answer. Every day, the student has to go through the deck trying to answer the question on the card and checking his answer. If the answer is correct the card will be repeated the next day, and if incorrect it will be repeated the same day. The main disadvantage of this system is that it becomes time-intensive when more flashcards are introduced, because the student has to go through all of the cards every day. Because of this, newer systems relying on spaced repetition were introduced, with which the time intervals between repetitions increase every time the student answers correctly Edge, Fitchett, Whitney, and Landay (2012). It was found that systems using adaptive algorithms was more effective and more satisfactory to the user than the other strategies (Edge et al., 2012).

The effects of flashcards have mainly been attributed to the spacing effect (Nakata, 2011; Edge et al., 2012), which means that repeated items are better remembered when both occurrences are separated by other events or items than when they are presented in immediate succession (Verkoeijen & Delaney, 2008; Logan, Castel, Haber, & Viehman, 2012; Siegel & Kahana, 2014; Xue et al., 2011; Karpicke & Blunt, 2011).

Concept maps

According to Eppler (2006), concept maps are defined as hierarchical graphs showing the relationships between concepts, including cross connections among concepts and their manifestations. They also define them by differentiating them from other available visual mapping techniques - mind maps, conceptual diagrams and visual metaphors - by several factors. The function of a concept map is to show systematic relationships among sub-concepts relating to one main concept, without the need for a proven analytic framework or picture. This makes it easy to apply in classroom teaching and self study and a useful support tool for summarising key course topics. The main topic should be placed at the top, so the hierarchy of the concepts is conveyed intuitively. All concepts are represented by boxes or bubbles with text, and the relations by labeled arrows. It is a flexible tool and has a high understandability by others compared to the other techniques. It however does not necessarily assist memorability, and it is not easy to apply by novices and requires extensive training.

Flashmaps

The flashmap is intended as an integration between the flashcard system and the concept map. The system uses a predefined concept map constructed by an expert which the users have to rote memorise. The concept map has the best combination of understandability and extensibility, whereas the low memorability is compensated by the flashcard system, and because it is predefined the construction difficulty is taken away. Where a flashcard system would then show a question, the flashmap shows a part of this concept map with one or more empty. The user then has to think of which concepts would fit in these nodes, and when requesting the answer the flashmaps shows the actual concepts. The student then can indicate per node whether he was right or not, which will then be used for rescheduling the nodes according to the adaptive algorithm (see figure 2). Canas and Novak (2012) describes that fill-in-the-cmap or memorise the concept map conditions are not recommended, because of the information in memorised concept maps not being integrated with other relevant knowledge and the lack of learners being actively engaged in assimilating new concepts and propositions into their cognitive structures. However, they do not provide statistics or literature in order to support this claim, and furthermore the findings from Karpicke and Blunt (2011) about paired associate learning being more effective for meaningful learning than concept mapping puts this claim into doubt.

Finally, the flashmap creates the opportunity for a more interactive concept map that starts with a parsimonious and theme-oriented structure which gradually expand the details along with the instruction, which is hypothesised by Tzeng (2010) to mitigate map shock. This phenomenon occurs when users view the kind of larger concept maps that might more fully capture textbook knowledge structures, but is a type of cognitive overload that prevents students from effectively

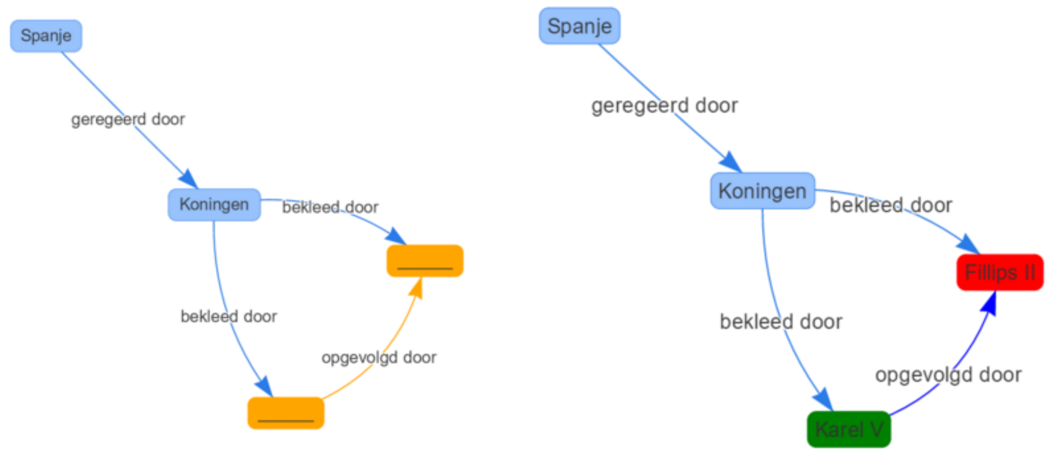

Figure 2: A display of the flashmap system, where the user has to think of the concepts fitting in the orange nodes on the left, and has to indicate which nodes were correct on the right

processing the concept map and thereby inhibiting their ability to learn from it (Moore, North, Johri, & Paretti, 2015). This mitigation will be facilitated by scheduling the central concepts towards the beginning and the details towards the end.

Research Question and Model

For researching the effects of the flashmap system relative to the effects of the flashcard system, it is important to consider two main factors: its actual benefits (research question **??**a and b), and its perceived benefits (research question IIa and b). Furthermore, for the validity of the system and of the experiment it is important to investigate how the system was used by the students (research question III).

To research whether the flashmap system is more effective or efficient than the flashcard system, the learning gain of high school the students will be measured, referring to the knowledge obtained by a student over the course of an instruction. Sequentially, the efficiency of the system is determined by the learning gain controlled for time spend on the system.

For measuring the affectiveness of the systems, the Technology Acceptance Model by Davis, Bagozzi, and Warshaw (1989) will be used (see figure 3). This model predicts the use of an information system by measuring the Perceived Usefulness and the Perceived Ease of Use of the user. These variables are mediators between External Variables and Attitude toward using, leading to Behavioural intention to use, which in turn leads to the Actual system use.

Finally, for the answering the final question an interview will be conducted with a sample of the participants, and by the server logging usage information about the user.

This leads to the following research questions: Regarding high school students learning for Dutch literature using the flashmap system in comparison to them using the flashcard system...

Ia. ...is the learning gain larger?

Ib. ...is the learning gain larger controlled for the time spend with the system?

IIa. ...do they perceive the system to be more useful?

IIb. ...do they perceive the system to be easier to use?

III How did the students use the flashmap or flashcard system?

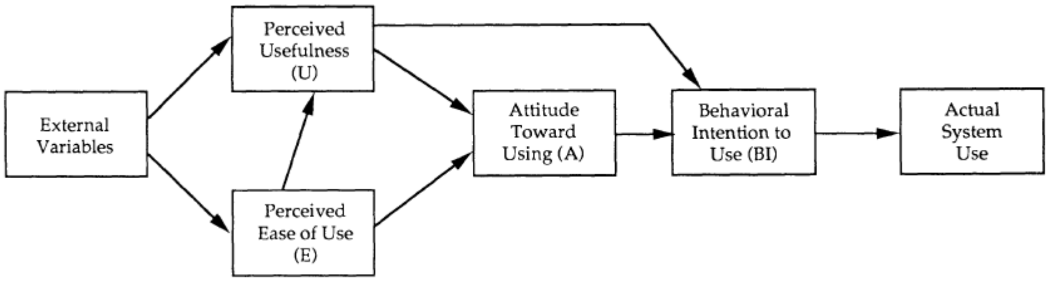

Figure 3: The Technology Acceptance Model by Davis et al. (1989)

Scientific and Practical Relevance

Answering the research questions has both practical and scientific relevance. From a practical perspective, it has potential to overcome the criticism from various authors about flashcard systems and answer the need for meaningful rote memorisation. From a scientific perspective, it could confirm the hypothesis by Tzeng (2010) that an expanding concept map might mitigate map shock. It also makes way for new research opportunities, for example what the effect is of integrating the flashmap with the games condition formulated by Canas and Novak (2012).

Research Design and Methods

Research design

Research questions Ia, b, IIa and b will be investigated using intervention-based research. Because of the systems being used for self-study by the students, they can be individually assigned to a condition, and this enables the use of a true experimental design. Since this will provide the most valid and reliable results, this research design is implemented in this experiment.

Furthermore, research question III is a qualitative research question, and therefore an interview will take place after the experiment in order to investigate how the students used the system. Next to the interviews, user data and actions will be logged by the server.

The quantitative and qualitative results will be mixed for the purposes of triangulation and expansion as described by Schoonenboom (2014). The interviews and logs could provide insight in the degree of which the systems were used the intended way and in why students had certain perceptions on using the systems. Both triangulation and expansion will be on a partial level of mixing, will take place concurrently, and the quantitative data will be dominant, since the qualitative data exists only to triangulate and expand the quantitative data.

Respondents

100 15 to 17 year old tenth grade Dutch high school students will be approached. They already have to prepare themselves for an exam on the same topic and thereby have incentive to learn. To increase the response rate, the students will be rewarded with a € 5 voucher for participation. The participants will be assigned to either the flashcard or the flashmap condition at random.

Instrumentation

The learning gain will be measured by the means of a pre- and post-test. Both tests will consist of random items from an item bank measuring both knowledge and comprehension levels of the students (Bloom et al., 1956). The tests will be directly based on the concept map, and also will be evaluated by the teacher in order to increase its validity. By using an item bank, the tests will be comparable and thereby the learning gain can be determined by subtracting the score on the pre-test from that on the post-test. Finally, the controlled learning gain is calculated by dividing the learning gain by time spent on the software. The survey will be an adaptation of the standardised Technology Acceptance Model questionnaire Davis (1989).

The interviews will be conducted using a topic list (Baarda, de Goede, & Teunissen, 2009), including "frequency", "usefulness", "ease of use", "external conditions", and "attitudes", based on the Technology Acceptance Model. The server logs will contain information about the reaction times, the correct responses, the nodes studied, the time investment, the IP address, and the client, which will be registered per user and per session.

Procedure

An outline of the procedure is given in figure 4.

Before the experiment takes place the experiment will have to be approved by the ethics committee from the University of Twente. On approval, the students and their parents will be briefed by means of a letter, which consists out of a general description, conditions (voluntary participation and withdrawal at all times), and rewards. They will also both be asked to fill in an informed consent form.

After that, the students with consent will be provided with a general introduction on flashcards by both the teacher and the researcher within the classroom. Then, when the students log into the system for the first time, the server assigns them randomly to either the flashcard or the flashmap condition. By making the introduction ambiguous enough, the students will not be able to recognise this condition in order to guarantee a double-blind experiment.

Before they start using the system they will be asked for general descriptive information such as date of birth and gender. A code will be assigned to them making it only able for the teacher to determine their identity. After that the pre-test follows, and for the next week will use the system daily for fifteen minutes. Finally the post-test and survey will be conducted. At the end of the post-test, the students can also indicate whether they are willing to participate in the interview.

Data Analysis

Research questions Ia, b, IIa, and b will be assessed by means of a t-test, using the following hypotheses:

Ia	H0:	$LG_{fc} \leq LG_{fm}$
	Ha:	$LG_{fc} < LG_{fm}$
Ib	H0:	$LGC(fg) \leq LG_{fm}$
	Ha:	$LGC_{fc} < LG_{fm}$
IIa	H0:	$U_{fc} \leq U_{fm}$
	Ha:	$U_{fc} < U_{fm}$
IIb	H0:	$E_{fc} \leq E_{fm}$
	Ha:	$E_{fc} < E_{fm}$

where LG = learning gain, LGC = controlled learning gain, U = perceived usefulness, E = perceived ease of use, fc = flashcard condition and fm = flashmap condition. For determining the learning gain, the pre- and post-test have to be scored with a predetermined rubric. The answers will be scored without the scorer being aware whether the question was asked within the pre- or the post-test, or which participant filled in the answer. After both the teacher and the researcher have scored a sample of the answers, the inter-rater reliability will be calculated. The rest of the answers will be scored by the researcher only, and after scoring all of the items the reliability of test items will be assessed further using Item Response Theory (?, ?).

The interviews will be transcribed and coded according to Baarda et al. (2009), and another inter-rater reliability will be determined by a sample of the interviews coded by the researcher and a peer researcher. The coded fragments will be checked to validate the results from the t-tests, together with the server logs made during the experiment.

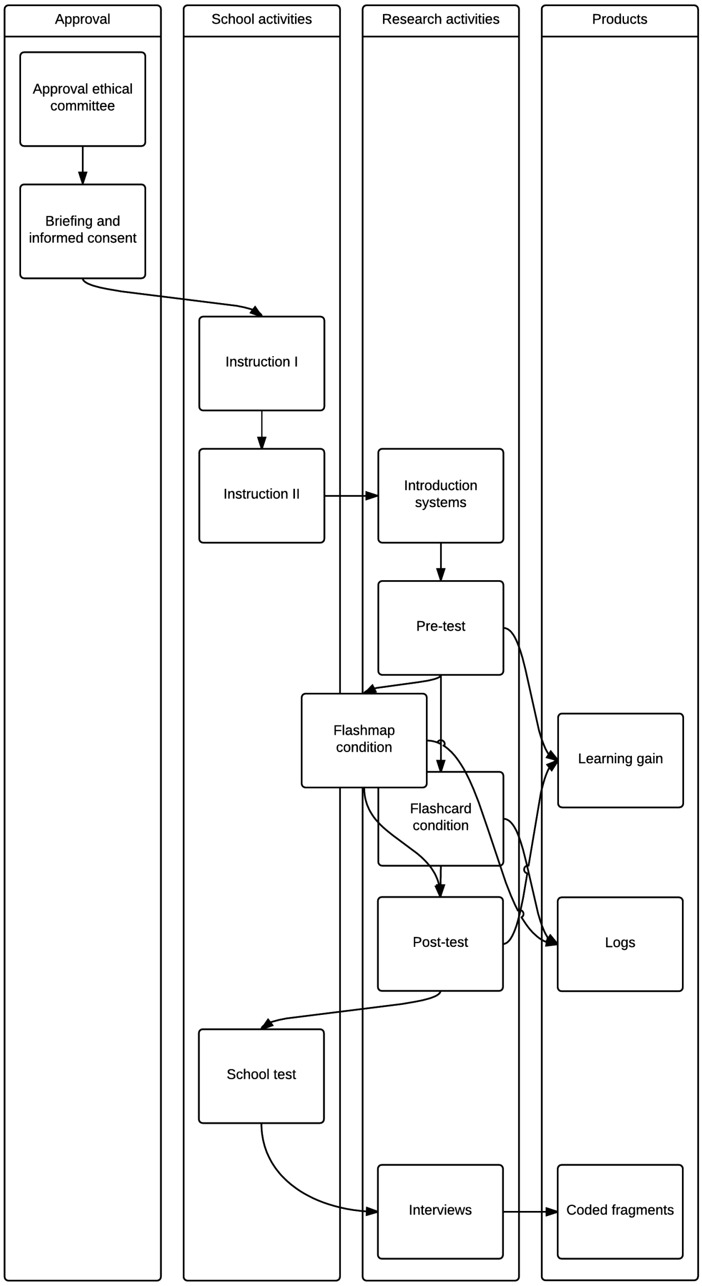

Figure 4: An overview of the different steps conducted within the research procedure

Planning

Timeline

Table 1 displays the activities together with the intended week number of execution and generated outputs.

Outputs

Their are several products being developed and data being collected during the project. First of all, a concept map will be developed containing the main concepts on 17th century Dutch literature, which will be the basis for the flashcards and the item bank questions. Furthermore, software will be developed being capable of both providing flashcards and flashmaps. The data being collected includes initial and final student answers, questionnaire results audio conversations, and server logs from the interviews, which will then be used for conducting the t-tests and triangulation and expansion purposes. This will finally be documented in the research report.

Week №	**Activity**	**Outputs**
16	Request approval ethical committee	Approval ethical committee
17	Finalising concept map	Concept map
17	Finalising flashcards	Flash cards
17	Finalising item bank questions	Item bank questions
18	Finalising adaption Technology Acceptance Model questionnaire	Adapted Technology Acceptance Model questionnaire
18	Evaluation concept map with teacher	Improved concept map
18	Evaluation flashcards with teacher	Improved flashcards
18	Evaluation item bank questions with teacher	Improved item bank questions
19	Finalising software	Flashcard and Flashmap software
19	Briefing and informed concent	Approval parents and children
19	Instruction I	
20	Instruction II	
20	Introduction flashcard and -map systems to students	
20	Pre-test	Initial answers
20	Experiment	User data
21	Post-test	Final answers
21	Questionnaire	Perceived usage and ease of use
21	Finalise scoring sample items	Scored sample items
22	School test	
23	Interviews	Audio conversations
23	Determine inter-rater reliability with teacher	Cohen's kappa
25	Finalise scoring all items	Scored items
27	Finalise transcribing interviews	Transcribed interviews
28	Finalise coding interviews	Coded fragments
29	Perform t-tests	Quantitative results
32	Finalising report	Research report

Table 1: The intended planning for research activities and outputs during this project

References

Anderson, J. (1982). Acquisition of cognitive skill. *Psychological Review*, *89*(4), 369–406. doi: 10.1037/0033-295X.89.4.369

Baarda, D., de Goede, M., & Teunissen, J. (2009). *Basisboek kwalitatief onderzoek: handleiding voor het opzetten en uitvoeren van kwalitatief onderzoek.* Groningen/Houten: Noordhoff Uitgevers.

Bloom, B., Engelhart, M., Furst, E., Hill, W., & Krathwohl, D. (1956). *Taxonomy of educational objectives: The classification of educational goals. handbook i: Cognitive domain.* New York: David McKay Company.

Burgess, S., & Murray, A. (2014). Use of traditional and smartphone app flashcards in an introductory psychology class. *Journal of Instructional Pedagogies*, *13*, 1–7.

Canas, A., & Novak, J. (2012). *Freedom vs restriction of content and structure during concept mapping: possibilities and limitations for construction and assessment.*

Davis, F. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319–340. doi: 10.2307/249008

Davis, F., Bagozzi, R., & Warshaw, P. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, *35*(8), 982–1003. doi: 10.1287/mnsc.35.8.982

Edge, D., Fitchett, S., Whitney, M., & Landay, J. (2012). *Memreflex: Adaptive flashcards for mobile microlearning.* doi: 10.1145/2371574.2371641

Eppler, M. (2006). A comparison between concept maps, mind maps, conceptual diagrams, and visual metaphors as complementary tools for knowledge construction and sharing. *Information Visualization*, *5*, 202–210. doi: 10.1057/palgrave.ivs.9500131

Ericsson, K., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, *102*(2), 211–245. doi: 10.1037/0033-295X.102.2.211

Golding, J., Wasarhaley, N., & Fletcher, B. (2012). The use of flashcards in an introduction to psychology class. *Teaching of Psychology*, *39*(3), 199–202. doi: 10.1177/0098628312450436

Hulstijn, J. (2001). *Intentional and incidental second-language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity.* Cambridge University Press.

Hwang, G., Wu, P., & Ke, H. (2011). An interactive concept map approach to supporting mobile learning activities for natural science courses. *Computers and Education*, *57*, 2272–2280. doi: 10.1016/j.compedu.2011.06.011

iMcCullough, C. (1955). Flash cards - the opiate of the reading program? *Elementary English*, *32*(6), 39–381.

Joseph, L., Eveleigh, E., Konrad, M., Neef, N., & Volpe, R. (2012). Comparison of the efficiency of two flashcard drill methods on children's reading performance. *Journal of Applied School Psychology*, *28*(4), 317–337. doi: 10.1080/15377903.2012.669742

Karpicke, J. (2012). Retrieval-based learning: Active retrieval promotes meaningful learning. *Current Directions in Psychological Science*, *21*(3), 157–163. doi: 10.1177/

0963721412443552

Karpicke, J., & Blunt, J. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*, *331*, 772–775. doi: 10.1126/science.1199327

Kornell, N., & Bjork, R. (2008). Optimising self-regulated study: The benefits - and costs - of dropping flashcards. *Memory*, *16*(2), 125–136.

Krathwohl, D. (2002). A revision of bloom's taxonomy: an overview. *Theory into practice*, *41*(4), 212–264.

Logan, J., Castel, A., Haber, S., & Viehman, E. (2012). Metacognition and the spacing effect: the role of repetition, feedback, and instruction on judgments of learning for massed and spaced rehearsal. *Metacognition Learning*, *7*, 175–195. doi: 10.1007/s11409-012-9090-3

Moore, J., North, C., Johri, A., & Paretti, M. (2015). Effectiveness of adaptive concept maps for promoting conceptual understanding: Findings from a design-based case study of a learner-centered tool. *Advances in Engineering Education*, *4*(4), 1–35.

Nakata, T. (2011). Computer-assisted second language vocabulary learning in a paired-associate paradigm: a critical investigation of flashcard software. *Computer Assisted Language Learning*, *24*(1), 17–38. doi: 10.1080/09588221.2010.520675

Nesbit, J., & Adesope, O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research*, *76*(3), 413–448. doi: 10.3102/00346543076003413

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

Schoonenboom, J. (2014). Points of addition and points of integration: Why oine point of interface is not enough. , 1–13.

Siegel, L. L., & Kahana, M. J. (2014). A retrieved context account of spacing and repetition effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(3), 755–764.

Singh, I., & Moono, K. (2015). The effect of using concept maps on student achievement in selected topics in chemistry at tertiary level. *Journal of Education and Practice*, *6*(15), 106–116.

Subramaniam, K., & Esprivalo Harrell, P. (2015). An analysis of prospective teachers' knowledge for constructing concept maps. *Educational research*, *57*(3), 217–236.

Tzeng, J. (2010). Designs of concept maps and their impacts on readers' perfomance in memory and reasoning while reading. *JOurnal of Research in Reading*, *33*(2), 128–147. doi: 10.1111/j.1467-9817.2009.01404.x

Verkoeijen, P., & Delaney, P. (2008). Rote rehearsal and spacing effects in the free recall of pure and mixed lists. *Journal of Memory and Language*, *58*, 35–47. doi: 10.1016/j.jml.2007.07.006

von Glaserfield, E. (2001). Radical constructivism and teaching. *Prospects*, *31*(2), 161–173. doi: 10.1007/BF03220058

Xue, G., Mei, L., Chen, C., Lu, Z.-L., Poldrack, R., & Dong, Q. (2011). Spaced learning enhances subsequent recognition memory by reducing neural repetition suppression. *Journal of Cognitive Neuroscience*, *23*(7), 1624–1633.

Zirkle, D., & Ellis, A. (2010). Effects of spaced repteition on long-term map knowledge recall. *Journal of Geography*, *109*(5), 201–206. doi: 10.1080/00221341.2010.504780