Linköping University | Department of Computer and Information Science

Master thesis, 30 ECTS | Cognitive Science

2018 | LIU-IDA/KOGVET-A--18/002--SE

The Potential of Argument Mapping as a Tool for Teaching Critical Thinking in Secondary School

Potentialen av argument mapping som ett verktyg för att undervisa kritiskt tänkande i högstadiet

Tim Lidåker

Supervisor : Annika Silvervarg
Examiner : Arne Jönsson

LINKÖPING UNIVERSITY

Linköpings universitet
SE–581 83 Linköping
+46 13 28 10 00 , www.liu.se

Abstract

English

With our modern deluge of information through social media, news and blogs, the importance of giving our children a better understanding of source criticism and critical thinking has become increasingly clear. In the field of critical thinking, the use of argument mapping software, a visual way of structuring arguments, has been shown to increase the understanding of arguments, and college students using the method score as much as three times higher on critical thinking tests.

This thesis presents a simplified digital argument mapping tool, developed in order to explore the feasibility of argument mapping for students in secondary school (aged 13-15). The tool is then tested on two secondary school classes. The thesis shows that the students are proficient at basic argument mapping with the tool, but also that the more complex facets of the method require more instruction in order to be used by the students. Recommendations are presented for further development of the argument mapping tool, as well as further studies of critical thinking using argument mapping in secondary school. The thesis has shown the argument mapping has potential for use in secondary school, and should be further studied.

Svenska

Med vårt moderna överflöd av information genom sociala media, nyheter och bloggar, så ökar vikten av att ge våra unga en bättre förståelse för källkritik och kritiskt tänkande. I fältet kritiskt tänkande har en metod som kallas argument mapping, en visuell metod för att strukturera argument, visats öka förståelsen av argument, samt vid användning studenter i universitet har metoden ökat resultaten på test i kritiskt tänkande upp till tre gånger.

Uppsatsen presenterar ett förenklat digitalt argument mapping-verktyg, utvecklat för att studera hur elever på högstadiet (13-15 år) kan använda sig av metoden, som sedan testas på två högstadieklasser. Resultatet visar att studenterna klarar av grundläggande argument mapping med verktyget, men att de mer komplexa delarna av metoden behöver vidare instruktion för att eleverna ska kunna förstå. Rekommendationer presenteras baserat på den data som samlats in, både för vidareutveckling av det designade verktyget, samt för framtida studier av kritiskt tänkande och argument mapping i högstadiet. Uppsatsen visar att argument mapping har potential för användning av högstadieelever, och bör därmed undersökas vidare.

Acknowledgments

I want to thank my supervisor Annika Silvervarg for the guidance and support throughout, Nils Dahlbäck for his sharp criticism and clear guidance, and Fredrik Stjernberg for some great insights in the early phases of the thesis. I also want to thank Stefan Elmström for allowing me to take time out of his teaching to test the prototype with his students. Lastly I have to thank Wiktor Strandquist for help with the database setup, because without your help I would have lost my mind, and possibly some hair. And to the one student who wrote in the questionnaire, "It was [expletive] hard. For [profanity] sake, do something better with your life, you [epithet].", I have taken your criticism to heart, and will do my best in order to make something better of myself.

Contents

List of Figures

1 | Introduction

Fake news, information overload and post-truth have become increasingly discussed in recent years. The deluge of information that we are bombarded with is hard to deal with, especially for children and teenagers. This, coupled with the fact that social media now conveys a lot of news to people, and often in a format that promotes clicks and narrative over facts and journalistic rigor, creates a skewed view of the world for many. Even at the university level, people tend to read news, political speeches and arguments as narratives, as Dwyer, Hogan & Stewart lamented in a 2012 study.

> [...] most students do not even acknowledge that the deliberations of an author within a text represents an argument and instead read it as if it were a story.
> (Dwyer, Hogan & Stewart, 2012, p. 220)

If the narrative flows well and it does not contradict previous beliefs, then little attention is spent on the structure of the argument, and often it is not even considered as an argument at all. The questions that then is left by the wayside are important questions like: What is opinion and what is fact? Who wrote this? And if one gets into more formal questions: Do the premises of the argument even support the conclusion? How strong are the premises, how relevant are they? Is the conclusion drawn a sound one?

One of the identified reasons for this is that without specific education, we do not have the structure nor the disposition for critical thinking (Halpern, 1998). In other words, we have a lack of an instinct to look at an argument and realize it needs to be analyzed, and lack the methods or tools with which to analyze them. One of the main solutions to this problem is increased education in critical thinking, that is designed to give people methods and a disposition to think more critically about the information that surrounds them. One of the main recommendations from a consensus report on critical thinking, called the Delphi Report (Facione, 1989), dealt with the importance of teaching critical thinking at a younger age:

> From early childhood people should be taught, for example, to reason, to seek relevant facts, to consider options, and to understand the views of others. It is neither impractical nor unreasonable to demand that the educational system teach young people the habits of mind which characterize the good critical thinker, reinforce those practices, and move students well down the path toward their attainment.
> (Facione, 1989, p. 15)

Whether or not critical thinking can be easily transferred across domains, there seems to be value in giving the structure of critical thinking and source criticism earlier on, in order to make it part of the automatic response to arguments. It does not seem unreasonable to

wish our children are taught to think critically, as recommended by the Delphi Report. This recommendation has been followed to some extent, at least in guidelines and expressed in values of schools, albeit not as much in actual classrooms quite yet. The Swedish National Agency for Education for example claim in their guidelines of "The schools basic values and mission" that:

> The students should be able to orient themselves in a complex reality, with a large information flow and rapid changes. Study skills and methods to approach and use new information therefore becomes important. It's also important that the students develop their ability to critically analyze fact and relations, and realize the consequences of different alternatives.
> (Swedish National Agency for Education, 2011, p. 3) [translated by the author]

Given this information dense world, and its demand for more critical thinking, as well as the scientific consensus claiming that it is not only possible, but in fact needed in lower ages, there exists a need for more research into critical thinking and its potential in schools. In this thesis the aim is to explore one possible way in which this can be achieved: through a well studied approach to teaching critical thinking called "argument mapping". Argument mapping has been studied to some extent in university students and up, with good results Dwyer, Hogan & Stewart, 2010, 2012, 2015; van Gelder, 2015; Twardy, 2004; Alvarez Ortiz, 2007), but no attempt at testing the method in school below university level has been done.

Therefore, the purpose of this work is to develop a simplified digital argument mapping tool, with some basic instruction in argument mapping. The tool will then be evaluated with students in secondary school, in order to identify to what extent students of this age can use the tool, and how well they grasp argument mapping. In the end, this will hopefully stand as a first foray into introducing argument mapping in younger ages, and with further research and development, show how it can be used in order to prepare students for a life of critical thinking. Therefore the purpose of the thesis is to:

1. Design and evaluate a digital argument mapping tool in terms of usability
2. Identify if students are able to understand and perform basic argument mapping, and look at what kinds of errors they make

The main goal is to design a simplified argument mapping tool, and with it explore if students can learn to use argument mapping as a method, as well as illuminating any potential of teaching critical thinking with the help of argument mapping in younger ages. This way, the thesis will be able to give some basis for the feasibility of argument mapping in lower ages, show a potential tool to use when teaching younger students be critical thinkers, and give some recommendations for further study.

First, the thesis will cover the theoretical underpinnings of critical thinking and argument mapping. After that the implemented tool will be presented, and it's design considerations discussed briefly. Then the evaluation of the tool will be explained, followed by the results. Finally, the discussion and conclusions will be presented, with recommendations for further study.

2 | Background

In order to make the following chapters easier to follow, the two main concepts are quickly summarized:

Critical Thinking (CT) is the approach and disposition we have to break down arguments, ideas or problems. CT both involves the structure of critical thought, as in how to break down arguments and news for example, as well as the disposition, as in the instinct to deal with sometimes quite complex thinking. The concept has many facets, and therefore a large report of the consensus of what CT really meant was conducted in the late 80's, called the Delphi Report, summarizing the most important features of CT (Facione, 1989).

Argument Mapping (AM) is a visual method and a structured approach to break down and analyze arguments in a hierarchical manner. It can be seen as something of a middle step between formal logic and natural language, allowing for a more structured approach than to natural language, while still keeping the form of the argument somewhat close to the natural language.

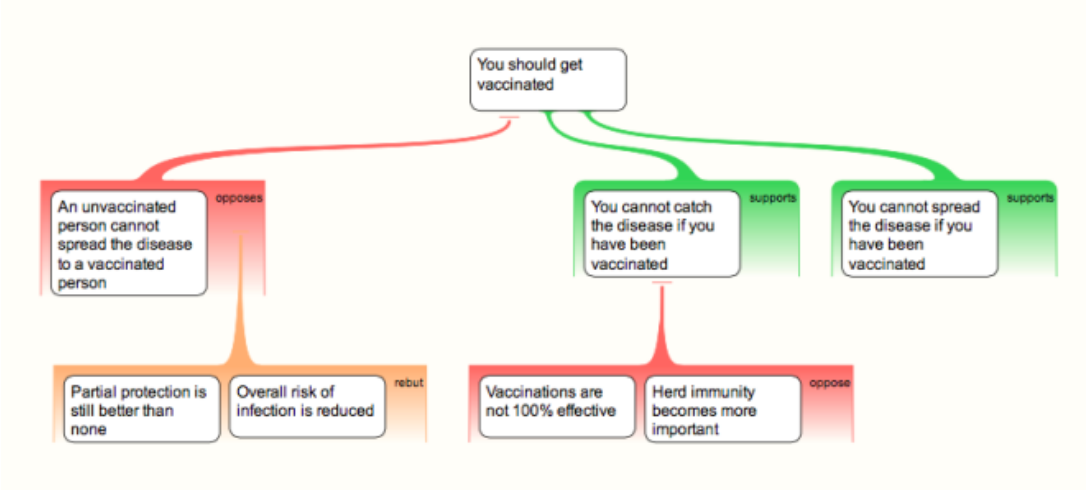

Figure 2.1: The Rationale AM tool

Example of an argument map using the Rationale software (Cavanagh, 2015)

AM has existed as a method for more than a century as a manual, pen and paper method (Whately, 1831), although it has gained significantly in popularity with the advent of software-based argument mapping. With these new kinds of software such as Rationale (van Gelder, 2002), creation, and more importantly revision of the maps have been made much easier, further playing into the strengths of AM.

The following sections will first present the defining traits and features of CT, and then deal with the specific method of AM and its uses within context of teaching CT.

2.1 Critical Thinking

The Delphi Report or "Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction" was a large study of the consensus in the field of CT, in order to create a clearer base of defining the concept (Facione, 1989). The report was a comprehensive summation of every facet of CT, according to a wide collection of American philosophy, education and psychology researchers, such as Ennis (1987), Lipman (1988), and Paul (1992). One of the main goals were to identify the skills and sub-skills that define a critical thinker, and thereby give a better basis for continued research in the field. The consensus list is a comprehensive step by step approach to finding, analyzing and evaluating arguments, which is the core of CT, according to the Delphi Report. A summarized version of the findings is called the "Consensus List of Cognitive Skills and Sub-skills".

CONSENSUS LIST OF CT COGNITIVE SKILLS AND SUB-SKILLS (Facione, 1989, p. 7)

SKILL	SUB-SKILLS
1. Interpretation	Categorization Decoding Significance Clarifying Meaning
2. Analysis	Examining Ideas Identifying Arguments Analyzing Arguments
3. Evaluation	Assessing Claims Assessing Arguments
4. Inference	Querying Evidence Conjecturing Alternatives Drawing Conclusions
5. Explanation	Stating Results Justifying Procedures Presenting Arguments
6. Self-Regulation	Self-examination Self-correction

What is identified as a central feature is the analysis, assessment and evaluation of arguments. As the report was done in the late 80's, AM-software had just started to appear, so the possibilities for breaking down arguments easily became more available than it had been before. The Delphi Reports further defines what features are important when analyzing arguments, stating that the following is important to identify and differentiate:

(a) the intended main conclusion,

(b) the premises and reasons advanced in support of the main conclusion,

(c) further premises and reasons advanced as backup or support for those premises and reasons intended as supporting the main conclusion,

(d) additional unexpressed elements of that reasoning, such as intermediary conclusions, unstated assumptions or presuppositions,

(e) the overall structure of the argument or intended chain of reasoning, and

(f) any items contained in the body of expressions being examined which are not intended to be taken as part of the reasoning being expressed or its intended background.
(Facione, 1989, p. 9)

These recommended points to focus on will be used as a foundation for how to design and evaluate the proposed argument mapping tool in this thesis. The features in terms of concrete AM is: Identifying the conclusion (a), identifying the supporting or negative premises (b), identifying support or opposition to premises (c), identifying irrelevant premises (d, f), and representing the intended overall argument (e). These are the main features that need to be learned in order to deal with arguments, in order to be able to effectively do argument maps.

Furthermore Halpern, another prominent CT researcher and creator of the HCTA test for evaluating CT ability (Halpern, 2018), has similar views on what is important structural aspects of CT, but instead comes from a more general approach to what a student needs to be educated in to become a good critical thinker.

(a) Verbal reasoning skills, in order to comprehend and defend against persuasive rhetoric

(b) Argument analysis skill, in order to better be able to see the structure of an argument

(c) Skills in thinking as hypothesis testing, in order to be able to weight possible explanations, thinking about sample size and so on. Essentially, acting like a scientist.

(d) Likelihood and uncertainty, in order to understand that things are not all or nothing, and that every premise and conclusion comes with different levels of certainty.

(e) Decision making and problem solving skills, in order to be able to generate and select between different solutions.
(Halpern, 1998, p. 452)

She begins with the verbal reasoning, and notes that it is regularly taught in rhetoric classes and in debate-focused CT, but follow it with the importance of argument analysis skills as a core element that needs to be understood before starting to assess likelihood, or doing hypothesis testing. This is where argument mappings steps in, and Halpern notes its usefulness in understanding arguments and its use in the teaching of CT.

2.1.1 Measurements of critical thinking

When it comes to measuring CT increase or decrease, there exists several standardized tests. Liu, Frankel & Roohr (2014) concisely summarize the most commonly used tests for pure CT in their paper Assessing Critical Thinking in Higher Education: Current State and Directions for Next-Generation Assessment.

> The majority of the assessments exclusively use selected-response items such as multiple-choice or Likert-type items(e.g., CAAP, CCTST, and WGCTA). EPP, HCTA, and CLA use a combination of multiple-choice and constructed-response items (though the essay is optional in EPP), and the Ennis–Weir test is an essay test. Given the limited testing time, only a small number of constructed-response items can typically be used in a given assessment.
> (Liu, Frankel & Roohr, 2014, p. 4)

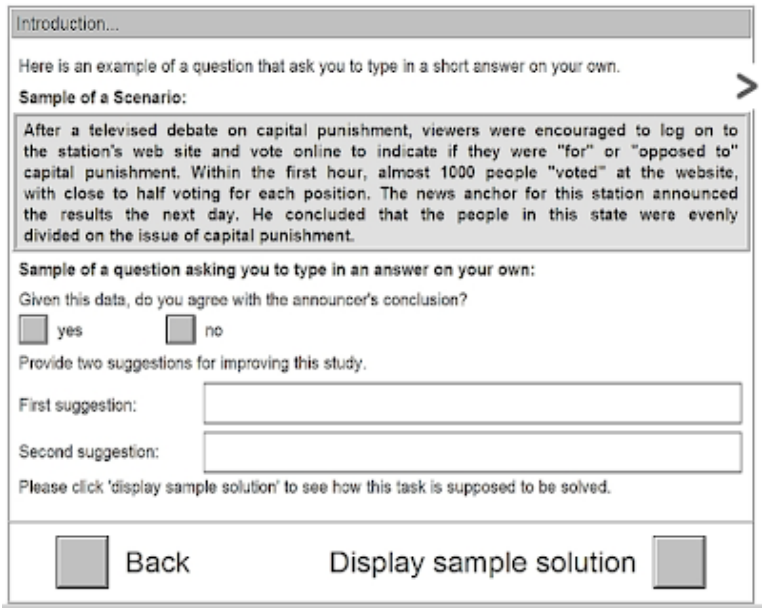

Figure 2.2: HCTA sample question

Sample HCTA question from Diane Halpern's blog (Halpern, 2018)

All of these tests still require a high proficiency in both English and CT, and is primarily designed for college students and up, as can be seen in the example question from the HCTA (2.2). The possibility of using these to evaluate younger, as well as non-native English speakers, is generally not advised. They are too difficult, and would likely need to be simplified greatly to be useful in evaluating grade schoolers. Many of these tests have however been used when testing AM efficacy, in order to have some baseline for the potential effects in teaching.

2.1.2 Methods for teaching critical thinking

One of the main sources of CT on university level is through philosophy, either through a philosophy major, or through individual courses meant as introductions to informal logic or CT (Alvarez Ortiz, 2007). With individual courses in some form of critical thinking the focus tends be more on the side of informal logic, with fallacies, syllogistic logic and biases, but may also have parts of formal logic, such as propositional logic, truth tables and predicate calculus. Alvarez Ortiz points out that the shared feature in most CT courses and general philosophical education is that there is extensive practice with arguments, and that this central to the teaching of CT, more than any one method is used as a standard.

One way of teaching younger children a kind of simplified CT is the P4C (Philosophy for children) program (Lipman, 1976). The program is in essence Socratic, a guided conversation about a problem or a concept, and revolves around questioning assumptions and trying to think collaboratively around problems. Although this is considered to promote philosophical thinking, which in turn promotes CT, the conclusion that the method increases CT is not definitive (Alvarez Ortiz, 2007).

When it comes to non-generalized instruction, there is some attempts at methods which focus on iterative, or decision tree-based reflection (Holmes, Wieman & Bonn, 2015). The focus with instruction in CT like this is in order for future researchers to be able to look at their

specific field critically, not in order to teach any type of general, day-to-day critical thinking. These methods are specifically designed for a high level of education, and require complex understanding of the material.

This leads us into argument mapping, one of the more well defined methods of teaching CT.

2.2 Argument Mapping

Argument mapping (AM) has seen increased use in recent years, but the full process of AM is not just drawing maps. In order to make the process of AM somewhat more clear Korb (2014) summarizes the steps of AM in a compressed way, from text to finished and evaluated argument, based on the argument analysis steps laid out in the book Reasoning by Michael Scriven (1979):

1. Clarifying Meanings - Clarify statements, and make them more direct, but try to capture the author's intent.
2. Identifying Propositions - Find the conclusions, premises, and sub-premises of the argument. Irrelevant rhetoric is stripped away.
3. Graph the Argument - Each premise is put as a node, and the relations between nodes are made as arrows. This is the main part of what makes AM different from other methods, the visual hierarchical structuring.
4. Make it Valid - Unstated or hidden premises should be added, in order to build a honest version of the argument.
5. Counterargue - Using the finished AM, identify weaknesses and strengths of the argument.
6. Consider Alternatives - Look at alternative explanations to weaknesses in the argument, or possible separate counterarguments. This is done not to "break" an argument, but to make sure all the relevant information is taken into consideration.
7. Evaluate - Look at the argument as a whole, and evaluate the validity of the argument. (Korb, 2014)

These steps are roughly representative of what AM is about, from text argument, to a finished and evaluated argument map. With the process a bit more clear, it's important to point out that the use of digital AM tools in order to teach CT is a somewhat new phenomenon, and while results have been very promising when looking at college aged students, no large scale attempts has been done looking at the efficacy for younger students.

2.2.1 Efficacy of argument mapping

Dwyer, Stewart and Hogan have explored AM's use in educational contexts, often comparing it to traditional CT education. When discussing the useful features of AM as compared to pure text they identify a few important aspects.

1. Unlike standard text, AMs represent arguments through dual modalities (visual-spatial/diagrammatic and verbal/propositional), thus facilitating the latent information processing capacity of individual learners.
2. AMs utilise Gestalt grouping principles that facilitate the organisation of information in working memory and long-term memory, which in turn facilitates ongoing CT processes.

3. AMs present information in a hierarchical manner which also facilitates the organisation of information in working memory and long term memory for purposes of enhancing and promoting CT.
(Dwyer, Hogan & Stewart, 2012, p. 12-13)

In this way, according to the authors, AM allows for structuring of an argument that makes it easier to assess and evaluate an argument. Sub-premises can be looked at separately, and their weight to the conclusion can be determined without having to keep the entire argument in memory.

Furthermore, in a 2012 study Dwyer, Hogan and Stewart showed a significant increase in CT skill using AM, measured with the HCTA, as compared to traditional CT courses (Dwyer, Hogan & Stewart, 2012). The study was done using e-learning software, and was a completely internet based course. While they showed increases in CT, they did not show any increase with the Motivated Strategies for Learning Questionnaire (MSLQ) or Need For Cognition Scale (NFCS). So while AM increased the CT results during post-testing, no increase in disposition or motivation could be measured by just using AM as compared to a non-AM approach.

Twardy (2004) compared traditional CT courses with courses using AM with the Reason!Able (now succeeded by Rationale) software (van Gelder, 2002), and found that even with teachers not skilled with the program or argument mapping would still result in the students showing a three times increase in score on the CCTST critical thinking test (Twardy, 2004). Testing with CCTST took place both before the course started to give a baseline, and after to measure potential increase in CT. In order to not only show that the students did better on the second test due to knowing how the test worked, they were compared with a control group that also did pre- and post-test, but did not receive any AM in their course.

Alvarez Ortiz (2007) found in a review of the empirical literature on CT, that CT courses using AM was about three times as effective at teaching CT as traditional, or what she calls Anglo-American analytic philosophy. Due to the quite clear advantage of using AM over traditional CT she closed with saying: "Philosophy departments are more effective than otherwise at teaching CT skills when it is done using argument mapping. Argument Mapping courses are by far the most effective way to improve CT skills." (Alvarez Ortiz, 2007)

2.2.2 Argument mapping and practice

All the previous mentioned studies that has shown increases in CT with the help of argument mapping has been in the context of a CT class, and it is possible that both the structure of the CT classes, as well as the time spent practicing has a large part in the increase. The classes then act as the scaffolding, and provides feedback that otherwise would be missing. This could be a possible obstacle if integrating AM into a non-CT based course. However, AM is argued to have specific features that increase learning of CT, and van Gelder (2015) hypothesized the following explanations of why AM in general works better than traditional CT:

1. that such software is more "usable" than the standard technologies we use for representing and manipulating reasoning;
2. that such software complements the strengths and weaknesses of our inbuilt cognitive machinery; and
3. that AM represents a semi formal "sweet spot" between natural language and formal logic.
(van Gelder, 2015, p. 190)

These hypotheses may be considered ”soft hypotheses”, as they are the result of observation and expertise gained over many studies involving AM, but not the result of any specific theorizing. This semi formal "sweet spot” may help when trying to break down argument, bringing them closer to our natural way of looking at arguments, while still giving structure. The fact is that we are very good at narrativizing arguments, but not very good at structuring them or breaking them down (Dwyer, Hogan & Stewart, 2012), yet students are seemingly much better at doing this kind of structuring CT when using AM, instead of just narrativizing.

Van Gelder (2002) is also a proponent of the LAMP (Lots of Argument Mapping) approach, that he developed when working with AM. He argues that it is necessary to do a lot of repeated AM, both to train yourself in the art of AM, but also to internalize the process. This way the students will be much more adept with AM, and in a sense have a greater disposition towards doing CT with AM when faced with an argument. The problem with the LAMP approach is that the amount of work required by a teacher becomes much larger with increased amounts of argument maps that need correcting and feedback to be written for the students.

In order to circumvent the otherwise large amount of time spent on correcting and providing feedback on AM problems, especially with a LAMP approach, Butchart et. al. (2009) attempted to create an automatic feedback system for a digital AM tool. They used the QP criteria in their design, but more importantly tried to make most of the feedback completely automatic, and provide hints of what was wrong and how to deal with it. The tool was tried in a 12-week course, with 8-take home exercises using the tool. The students were given both a pre- and post-test with the CCTST test. The result showed an increase in CT of about 0.45 standard deviations, which they claim is in the middle range of gains from AM when looking at the literature. They conclude that while the tool is limited in scope, and no replacement for an experienced teacher, it could be a way of offloading teachers when it comes to correcting and feedback to students. Still, they warn that any attempt at creating an automatic feedback system becomes increasingly difficult when dealing with larger arguments.

2.2.3 Appropriate age for argument mapping

While AM has not been studied in sub-university students, teaching of CT has been attempted in lower ages. As mentioned previously, the P4C (Philosophy for Children) (Lipman, 1976) was another attempt to teach children a kind of CT, through Socratic reasoning in a collaborative environment.

A similar kind of basic CT has been tested in lower ages, such as in the paper “Critical thinking in elementary school children” they look at the methods and practices appropriate for teaching CT in elementary school (Florea & Hurjui, 2015). The level of CT applied in elementary school aged students is more focused on a softer introduction to CT, and the paper summarizes the CT learned by the students as:

> [the] students gained speed in solving tasks, the effective data selection , have developed capacities to formulate arguments and opinions have proven autonomy thinking and problem solving logical paths showed multiple possibilities to solve and set the most appropriate solution for a given context.”
> (Florea & Hurjui, 2015, p. 571)

Given this it seems that the basics of CT through AM should be possible to teach in secondary school. Furthermore, looking to Piaget’s four stages of development (Inhelder & Piaget, 1958), the fourth stage the “formal operational stage” is where the abstract thought, metacognition, and problem solving really comes in, and has been shown to happen around age 12 . Given the metacognitive aspects of CT, and the problem solving involved with AM,

students over 12 should be good candidates for CT, especially through AM. This makes secondary school (ages 12-15 in Swedish schools) a reasonable place to attempt to use AM to learn CT.

3 | Design

The decision to create a new digital tool for argument mapping instead of using previously successful digital argument mapping tools is due to the complexity of these tools. Tools such as bCisive (Reasoning Lab, 2018), Rationale (van Gelder, 2002) and Argunet (Schneider, Voigt & Betz, 2007) are primarily designed for university students and up, and all have a host of complicated addons and features. This results in a steep learning curve, and learning the tools themselves being a significant part of education within argument mapping. As one purpose of this thesis is to evaluate how well students are able to learn the concept of argument mapping, and not their proficiency in a certain tool, these more complex tools were not considered candidates.

Instead a tool designed for simplicity was needed, where the interaction is simpler, and the features limited. What is needed is the ability to have a text shown, a direct way of transferring text to the argument map, and arrows indicating if a premise supports or do not support a conclusion.

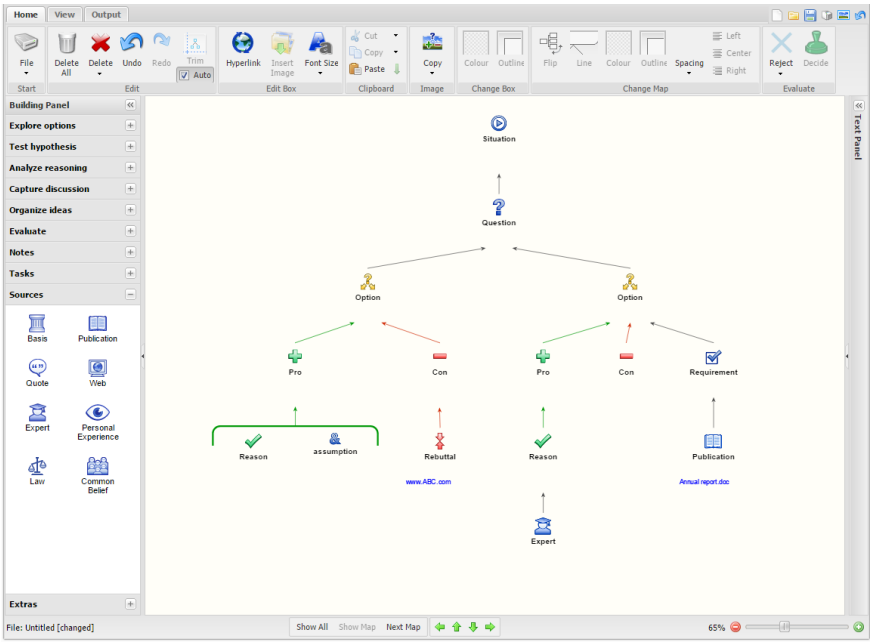

Figure 3.1: Interface of the AM tool bCisive

Showing the complexity of AM interfaces (Reasoning Lab, 2018)

3.1 Prototype

The prototype was developed using Javascript, HTML, PHP, and CSS, as well as the library GoJS (Northwoods Software, 2018) for Javascript. The choice of this implementation was in order for the tool to be instantly accessible (given internet connection), usable on every modern computer, and with little to no setup time, and in order to have data collection in the form of a database. The tool was designed to remove as many of the complexities of modern argument map software, as well as simplify certain aspects of AM.

The tool focuses on a subset of the steps of AM, as stated by Korb (2014), presenting a simplified version of "2. Identifying Propositions", followed by "3. Graph the Argument". The tool does not delve into any of the evaluative work that the later steps deal with. The prototype also simplifies "1. Clarifying Meanings", as the conversion from natural language requires quite a lot of training to begin with. The way the thesis tool simplifies these difficult parts is through a simplification to how the premises are stated, in that the arguments presented are expressed with each sentence being one premise. This is done in order to test the argument mapping abilities specifically, and exploration of the full method of AM is left to future studies. The finished prototype is shown in figure 3.2, with an example of a finished argument map.

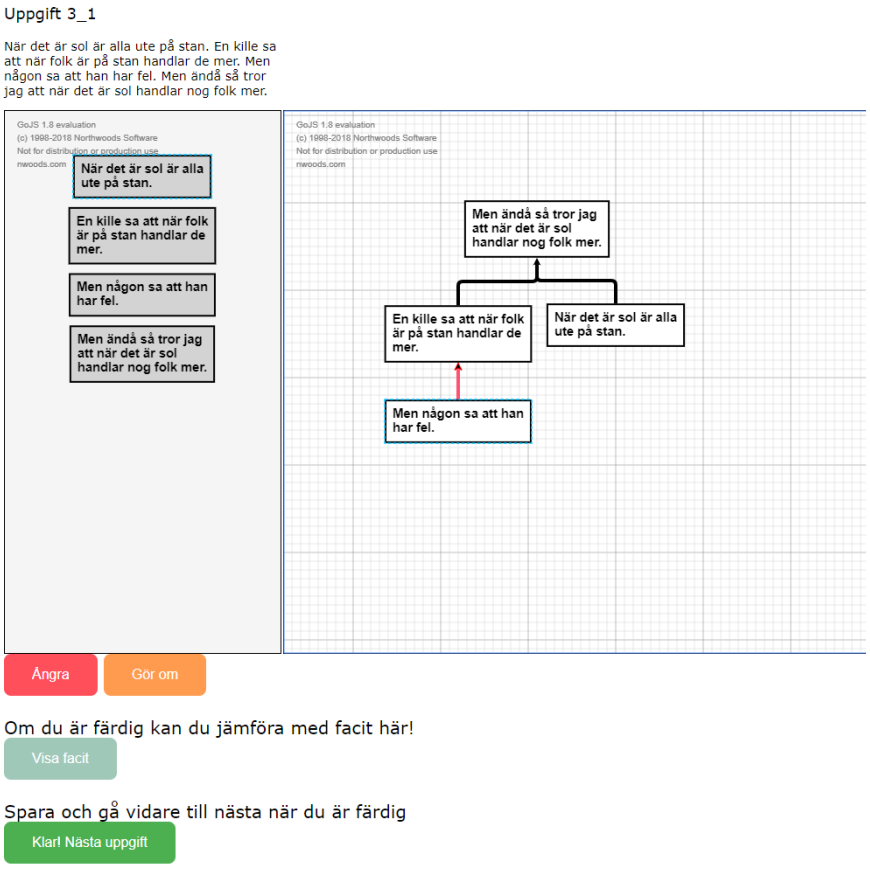

Figure 3.2: Interface of the thesis tool

An argument was shown at the top of the page, and then represented as separate boxes for each premise in the palette to the left. Each premise box could only be dragged onto the board once, in order to limit the chance of duplicate premises. The inclusion of the text form of the argument on top was in order for the students to first see the argument in its natural form, and read it as an argument and not just a set of disjointed statements. The boxes show the

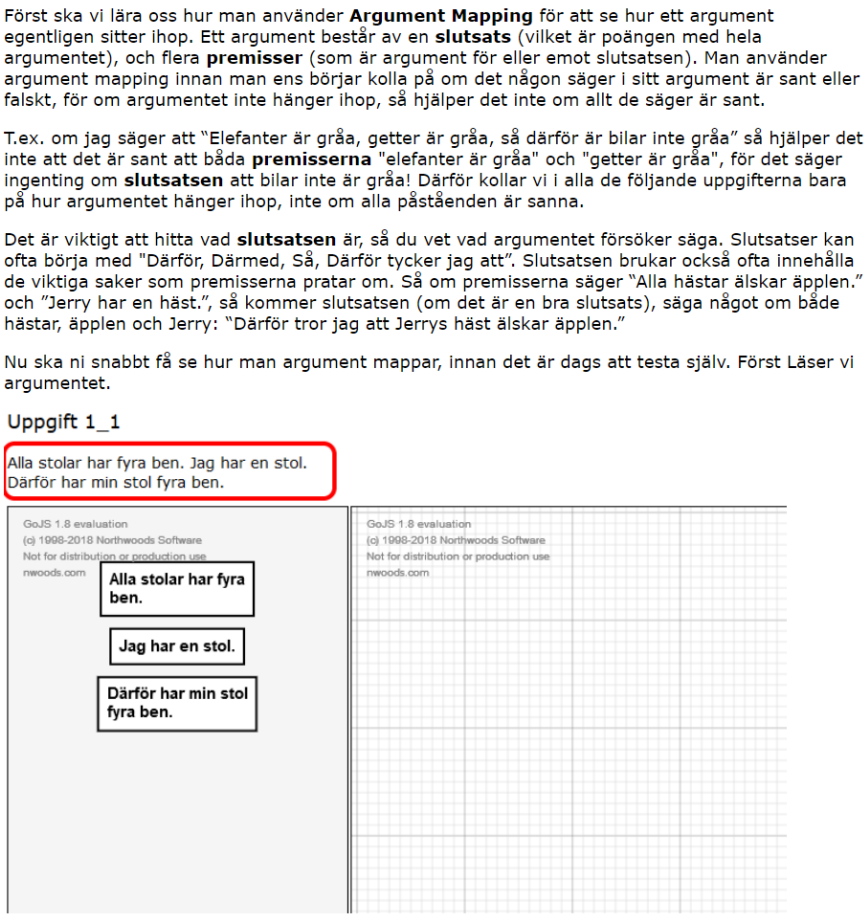

Figure 3.3: Short instructions for AM in thesis tool

Short introduction to the concept of argument mapping.
This is followed by further instruction in use of the tool.

argument separated into premises, without the students having to do the conversion from text to boxes by themselves. Again this is done to minimize general CT education for the evaluation, and allow the students to engage in the hierarchical mapping of premises as is central to the AM method.

These boxes could then be connected by dragging from the top of the box, to the bottom of another, creating an arrow link. This arrow could be double clicked to turn red, indicating a negative premise (this feature was introduced when negative premises were introduced, and could not be performed before). The link could be deleted by either double clicking, or pressing the "Undo" button. If an undo-action was incorrectly pressed by a student, they could press "Redo" to redo the action. The way the interaction with the arrows worked was made in order to make the user interface as minimal as possible, to reduce the time it takes for the students to learn the tool, so only one click would make a connection, and one click would make the arrow negative.

If the task was an introductory task, it had a "Show Correct Answer" button to show what a correct answer should look like, to give some feedback on the first few tasks. The tool, and each new feature was introduced with some text explaining it, as well as some images

Figure 3.4: Short introduction to irrelevant premises

to show the new features. In figure 5 and figure 6 the rough amount of instruction per new feature is shown.

3.2 Tasks

In addition, van Gelder (2001) argues that design of any digital AM tool should be according to Quality Practice (QP) in order to be successful at teaching students. QP is a hypothesis developed by van Gelder when looking at what features promotes learning, especially when it comes to CT. QP states that when it comes to complex problems such as is encountered in CT, the most successful way of learning how to deal with them is deliberate and repeated practice. The five criteria that QP recommends are: Motivated, Guided, Scaffolded, Graduated and Feedback (van Gelder, 2001). The tasks and instructions were all written in Swedish, as the testing occurred in Swedish schools. The QP criteria was used when designing the tool and it's tasks. As the tool is an early prototype, some concessions were made, but the criteria were implemented in the following ways:

- Motivated - the students deliberately practiced in order to deal with the tasks
- Guided - the students had instructions and clear delineation of problems
- Scaffolded - for each new feature introduced, there was three problems with correct solutions viewable
- Graduated - the tasks increased in complexity, both in the basic tasks as well as in the more complex arguments

- Feedback - again, there was three problems with correct solutions, and this was the only feedback the students were given.

When using the tool the students first got a short introduction to argument mapping basics and how to perform these actions in the tool. The first step, to identify the conclusion is explained, and then it is shown how you drag this box from the palette to the board. Then it is shown how you find the supporting premises, and how you connect them with arrows to the conclusion. With that explained, the students are allowed to start argument mapping.

There are four categories of tasks in the tool, done in succession. The categories are: "Basic argument mapping", "Irrelevant premises", "Negative premises", and "Complex argument". The first three categories consist of five tasks, of which three has a viewable correct solution. The final two in a set did not have a viewable correct solution, and are used to assess if the students has grasped the categories feature. Each category has slight scaling in difficulty, in order to see if there's a falloff in correct answers when the negative premise for example is less obviously referring to one premise.

The tasks were all written by the designer of the tool, as natural language syllogisms in the most basic tasks, and progressively adding irrelevant, negative, and more supporting premises. The tasks were not intended to all be valid arguments, but arguments expressed in roughly the same form as a syllogism. This was done in order to better capture natural arguments, and not only make toy examples which structure can be teased out without understanding the underlying relations. All tasks included in the tool, with English translations, can be found in Appendix B.

In the first category "Basic Argument Mapping" there is the basics of argument mapping with syllogisms, which is one fairly clear conclusion with two supporting premises (figure 3.5).

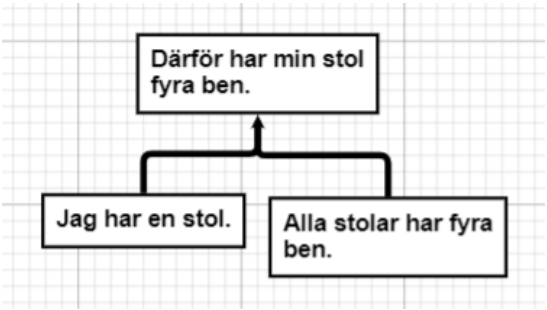

Figure 3.5: Basic AM

In the second category irrelevant premises are introduced, that is premises that neither support nor contradict the conclusion. The students are instructed to identify these, and put them on the board, but not connect them to the conclusion, as they do not support it (figure 3.6).

In the third category negative premises are introduced, which are premises that do not support, but instead contradict the conclusion (or possibly one of the premises). The students are shown how to make the connecting arrow red, in order to show that the premise is negative towards the connected premise (figure 3.7).

In the final category, complex arguments, there is ten tasks with increasing difficulty, both in complexity of the argument map as well as in amount of premises (figure 3.8). The first two of these had a viewable correct solution, but the following eight did not. In this category there

Figure 3.6: Irrelevant premise

could be one or more irrelevant and negative premises, in order to challenge the student to see if they could apply the previously learned skills in more opaque arguments. The size also increases with the scaling, with the first complex task having five separate premises, and the final one having eight separate premises.

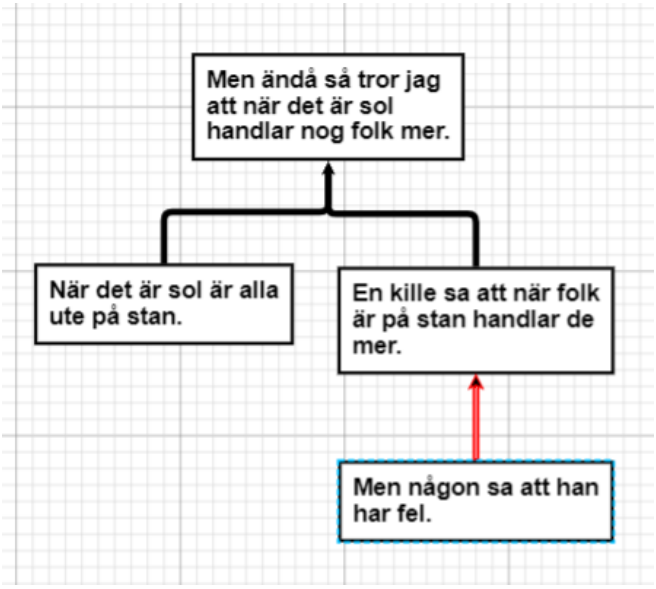

Figure 3.7: Negative premise, identified by the red arrow
(brighter grey arrow if viewing in black and white)

3.3 Pilot testing

The tool was pilot tested in order to be able to improve the tool, first with an adult individual, then with an appropriately aged individual (6th grade). The testing with the adult was done in order to give insight into what examples in the tool was strange, or where difficulty spiked too much, as well as to give general design input. The 6th grade student then tried a shortened version of the tool, with a reduced amount of basic training tasks, which were removed after the first testing session. The important flaws and bugs that were caught were the following:

- Too many simple training tasks, and too few complex ones
- The scaling of some tasks was too steep

- General comments on language and clarity in instruction
- The complex tasks had poor scaling, becoming too difficult too fast
- Presentation of the argument should be shown as a text, not just as a set of boxes, in order to promote students reading it as an argument, not just a set of disjointed statements

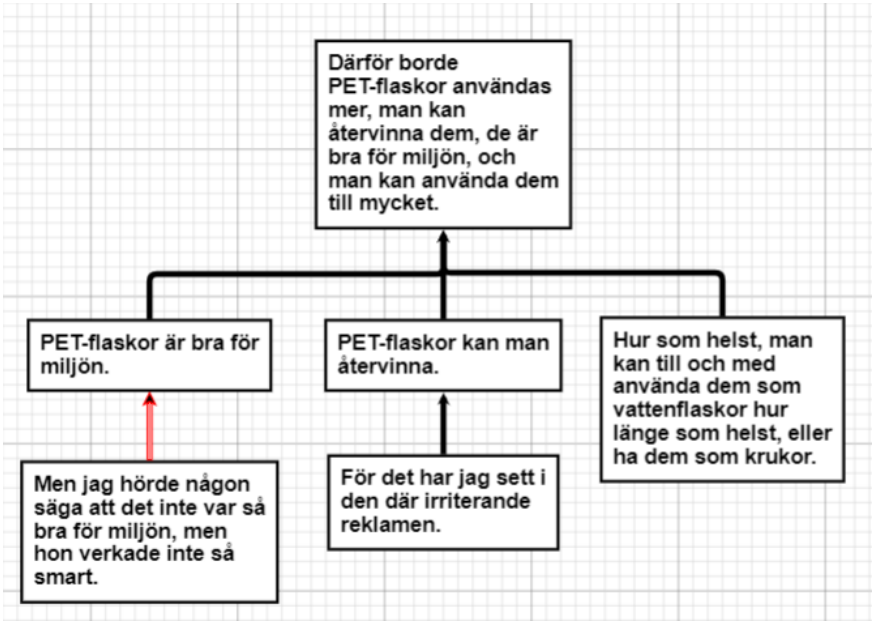

Figure 3.8: Complex argument

In addition to the design and scaling changes, the pilot testing showed that the basics of AM were communicated to an adequate degree to do the tasks successfully and the tool itself provided a good environment to do AM, being neither too complex nor hard to use. After the improvements the tool was tested in schools.

After the changes, the three Basic Tasks (Basic, Irrelevant and Negative premises) each consisted of three training examples with a viewable solution, and two test questions without a solution. The Complex Tasks scaling was adjusted, and had ten tasks, of which the two first had a viewable solution.

4 | Evaluation

In order to evaluate the designed tool, and the potential for AM use, testing in secondary school was conducted. The chapter will present the participants, the procedure of the testing, and the analysis of the data.

4.1 Participants

The classes tested was one 7th grade class (12-13 year olds) and an 8th grade class (13-14 year olds), resulting in 45 participants. The 7th grade class was split into two, and tested on two separate occasions due to their schedules. The 7th grade class was calm and silent, with all students working throughout and if done before the time was up, returned to their scheduled school work. The 8th grade was loud and less focused, and due to the teacher not being present, needed more supervision in order to stay focused. This could have some bearing on the result, and will be taken into account in the discussion.

4.2 Procedure

Testing of the tool was conducted by the designer of the tool, and some informal observation was done during testing, both for catching any bugs with the tool, and in order to identify problems with understanding AM when students asked for help, to give some context to the data. All testing and instruction was done in Swedish, including the tasks the students were to argument map. Before the testing was conducted, a passive consent form was sent to the students in order to describe the content of the study, how the data would be used, as well as informing the students and parents how to opt out of the study if they did not wish to participate.

Testing was conducted on three separate occasions, with testing of each class spread out over two consecutive days, with the student starting on the first day, and picking up where they left of on the second day. Following a short introduction of the purpose of the testing, the students were allowed to start using the tool. No extensive information about argument mapping was given in the introduction, but only referred to the fact that the tool would introduce the concept. The introduction also included a question as to if the students had ever heard about argument mapping. The students got to use the tool for 50 minutes before testing was stopped, and would continue for 50 minutes the following day.

Students went through the introductory tasks and eventually started on the complex tasks, and during the second day they picked up where they left of. If they managed to finish all the tasks before the time was up, either on the first or second day, they were asked to do their otherwise scheduled schoolwork until the time was up. Either when all the tasks were done, or when the second day was over, the students were asked to fill in a questionnaire.

The questionnaire contained a usability questionnaire (SUS) and a costume questionnaire with questions pertaining to the difficulty students felt with different aspects of argument

mapping. The SUS (System Usability Scale) questionnaire (Brooke, 1996) allows for a evaluation of the designed tool in terms of usability, and is a validated and well used tool (Bangor, Miller & Kortum, 2009). SUS is an effective way to classify the overall usability of a product, and in general is best used when comparing scores with other products, so it is not designed to be used to find specific problems with the product. In this thesis it is used mainly to establish that the designed tool is at least at baseline usable and will not stand in the way of learning AM in general.

The custom questionnaire was designed to identify perceived difficulty with AM, in order to identify any mismatches between the perceived difficulty and the actual results. It contained the following questions with a likert scale of 1-5:

- I found it hard to identify the conclusion
- I found it hard to know which conclusion/premise I was supposed to draw arrows to
- I found it hard to identify irrelevant premises
- I found it hard to identify negative premises

There was also two long text answers:

- Was there anything you got stuck on, or didn't understand?
- Did you at any point feel like you needed more information about argument mapping, or more help in order to finish a task?

If the student asked for help with understanding the tool, some clarifying tips would be given, such as how to delete a link that had been placed. No help was given if the question pertained to the argument mapping, and they were instead asked to do as well as they could with the information contained in the tool.

4.3 Analysis

Data was gathered by saving every finished task to a database as they were completed by the students. These were completely anonymized, and there was no way of seeing who made what task, or connect what one person made any two tasks. These finished tasks were then manually analyzed one by one, after criteria based on the Delphi Reports stated features of argument analysis (Facione, 1989): Identifying the conclusion (a), identifying the supporting or negative premises (b), identifying support or opposition to premises (c), identifying irrelevant premises (d, f), and representing the intended overall argument (e). This was done in order to identify and what kinds of errors were made, as well as if the students could grasp basic argument mapping. The criteria used in analysis the maps were:

Completely Correct	The AM has no errors
Completely Wrong	The AM is completely wrong
Conclusion Correct	The conclusion is correctly identified
Conclusion Wrong	The conclusion is incorrectly identified
Irrelevant Correct	Irrelevant premise identified
Irrelevant Wrong	Irrelevant premise incorrectly identified
Negative Correct	Negative premise identified
Negative Wrong	Negative premise incorrectly identified
Support Wrong	Support for conclusion or sub-premise incorrect
Incomplete	No boxes or no links at all

There is no classification of "Support Correct", as if the conclusion is correctly identified and the supports correct, the argument is Completely Correct, and if the conclusion is incorrectly identified the supports are incorrect. On the other hand, the support can be wrong, but you may still have identified the Correct Negative for example, but the inverse is not helpful to classify.

If the AM was not completely correct or completely wrong, first it would be determined if the conclusion was correctly identified, then if any possible irrelevant premises were correctly identified, followed by negative premises and finally the support for the conclusion.

The granularity of the criteria ignores to some extent different kinds of classes of mistakes. For example, a task would get the note "Negative Incorrect" both if the negative premise of the argument was not identified at all, and if a non-negative premise was identified as negative. The mistakes are different, but they both show a problem with identifying negative premises, and therefore falls under the same classification in this system. Figure 4.1-4.3 shows these different ways of constructing an AM that results in the classification.

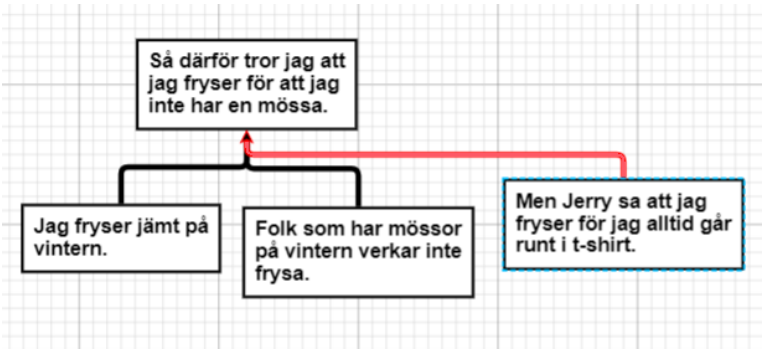

Figure 4.1: Completely Correct

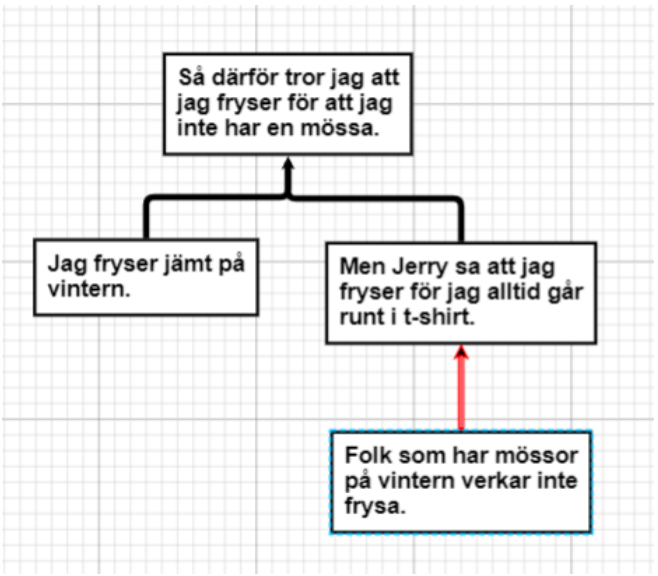

Figure 4.2: Incorrect negative (Classification: Incorrect Negative)

If a students AM had not given the intended interpretation, but had an interpretation that could be possible given the argument, the task would get a "Completely Correct" tag. Differing interpretations like this is common when doing AM, as the writer of the argument may have been somewhat unclear what a certain premise supports. Completely Wrong was given

if the map had completely misidentified conclusions and support, or was not an attempt at creating an AM at all, as seen in figure 4.4.

If a task had no arrows or boxes at all, or just boxes but no arrows, they were classified as Incomplete. These were assumed to be either due to a fault with the tool, or due to the student not actually trying to make an argument map (figure 4.5).

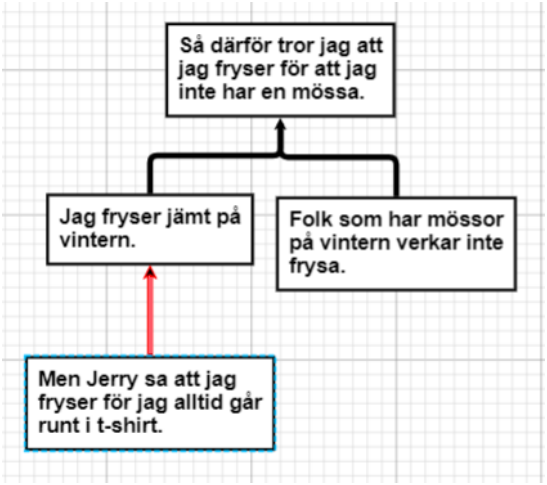

Figure 4.3: Correctly identified negative, but connected to the wrong premise (Classification: Incorrect Negative)

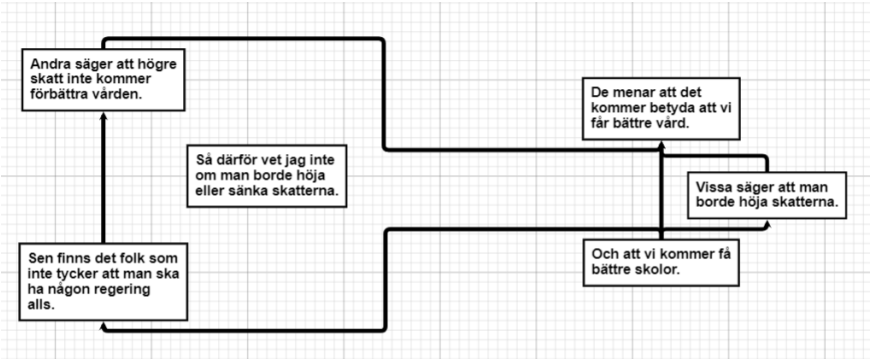

Figure 4.4: No clear structure (Classification: Completely Wrong)

Figure 4.5: No connections (Classification: Incomplete)

5 | Results

5.1 SUS

The full questionnaire, with SUS and a custom questionnaire, was filled out by 20 out of 20 of the 8th graders and 24 out of 25 of the 7th graders. The result of the SUS questionnaire gave a score of 63.5. The score is considered between "OK" (50) to "GOOD" (70) according to Bangor, Miller & Kortum (2009), with 70 being the average for a finished product. Specific question in the SUS were however lower than the average. The questions 6 (There is too much inconsistency in the tool) and 10 (I need to learn much before I can become productive with the tool) received lower scores compared to the other questions. 6 received a 1.75 average, and 10 received a 1.92 average. This indicates either difficulty with the tool, or possible difficulty understanding the questions, which was something that came up during the observation of the students.

5.2 Argument Mapping Tasks

All of the students in the 8th grade (20/20) completed all tasks, while in the 7th grade only 79.1% (19/24) complete all tasks. Both classes had the same amount of time to interact with the tool. The full table of tasks and results can be found in Appendix A.

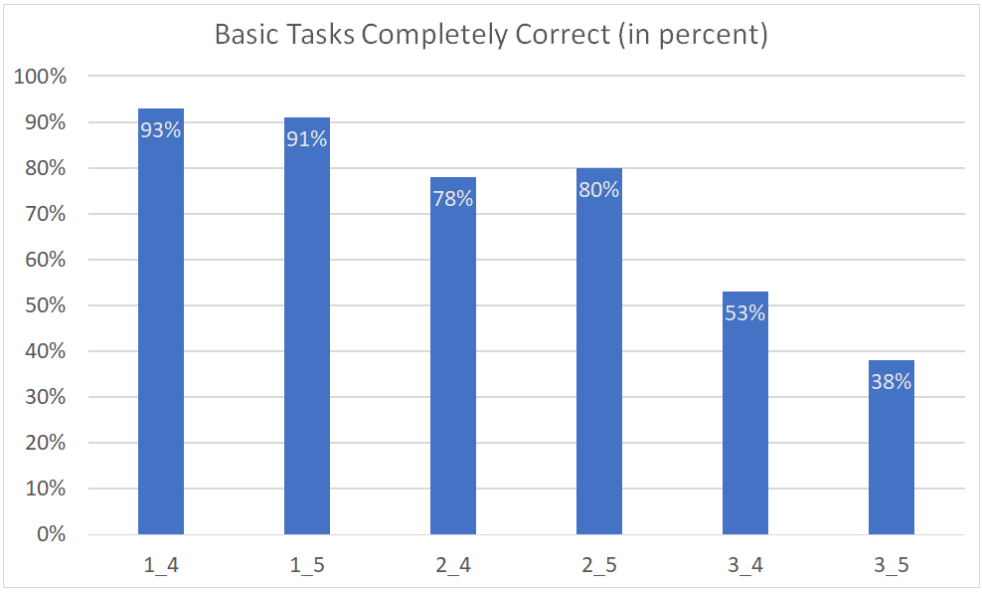

Figure 5.1: Basic tasks

In the basic tasks (as seen in figure 5.1), Basic Argument Mapping, Irrelevant Premises, and Negative Premises, there was a high amount of completely correct answers, falling off more

and more towards the end of the basic tasks. The negative premises provided the most challenge, with only 53% on the first task, and 38% on the second task were completely right.

When observing the two classes, four students in the 8th grade completed the full set of tasks within 20 minutes, with most answers being either incomplete or completely incorrect. This accounts for some of the lower amounts of correct answers, but most other students spent the full time allotted for the tasks. The negative premises had the least amount of completely correct answers, for both classes.

Of the tasks that were not completely correct, most of the errors occurred when introducing the negative premises, with 27 errors total. A majority of the incorrect argument maps were still finding the correct conclusion, but failed to identify or correctly connect the negative premise.

In the Complex Tasks, the students had difficulty in making the argument maps completely correct. The full amount of tasks done in this part was 325. As seen in figure 5.2, a the majority of the tasks, had the conclusion correct (65%), if not completely correct or completely wrong. The low occurrence of Conclusion Wrong (2%) is likely due to the fact that if the conclusion is wrong, the entire structure of the argument tends to be wrong. Therefore most of the misidentified conclusions get classified as completely wrong, due to support and other premises being restructured erroneously.

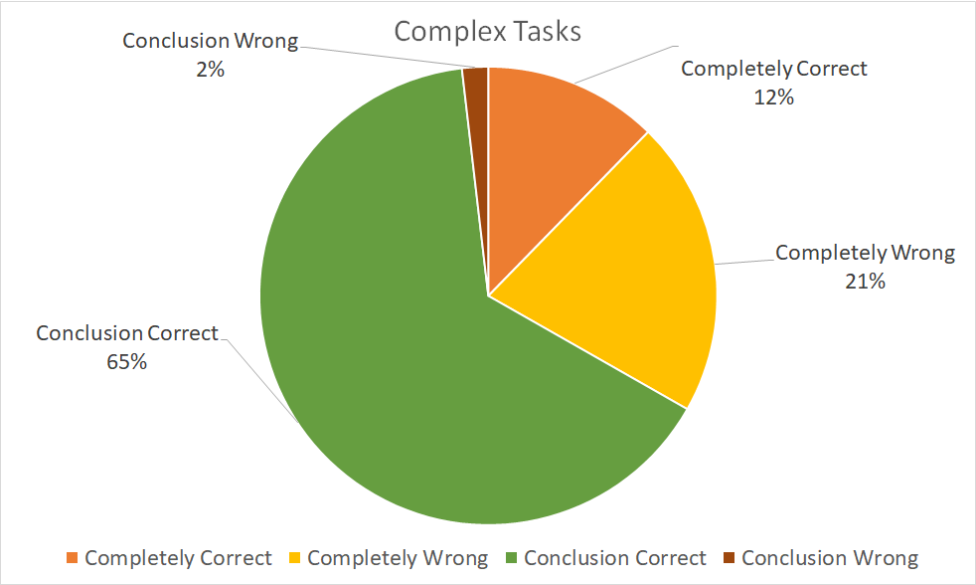

Figure 5.2: Distribution of correct and incorrect on complex tasks

If we look at the kinds of errors made by the students, we see where most errors occur. Given the possibility of making more than one mistake per task however, the amount of errors adds up to more than the amount of tasks.

Figure 5.3 shows that a large percentage (42%) is due to incorrect supporting premises. The errors in irrelevant or negative premises were roughly equal, at 28% for the irrelevant, and 30% for the negative.

A somewhat common mistake the students make is what has been classified as "Narrativizing", as seen in figure 5.4. This is when the argument has been connected more in line with the flow of the argument, and not in the forms of premises leading to conclusions. Then the map has premises that do not actually support another premise at all, they just occurred in

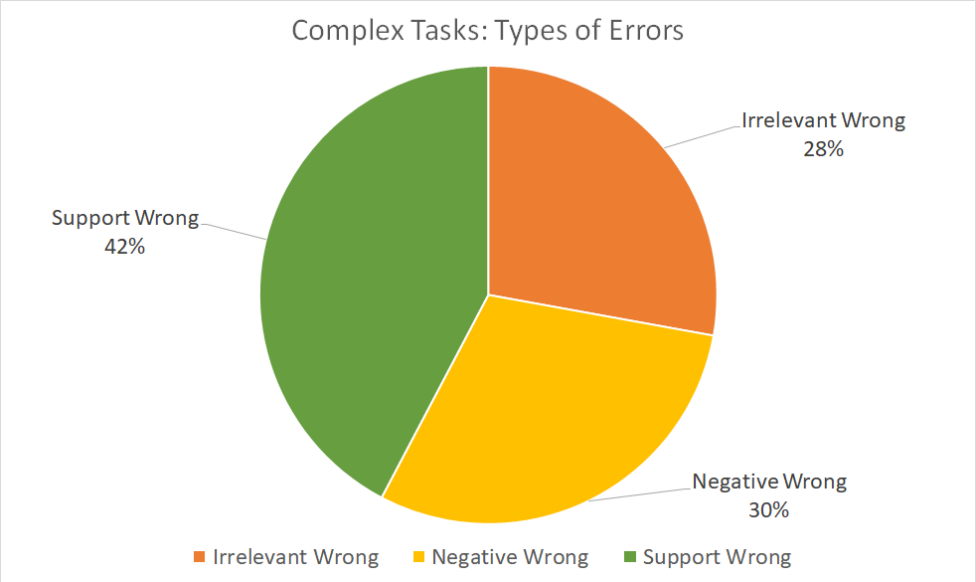

Figure 5.3: Distribution of kinds of errors on complex tasks

succession in the text. This is in line with what Dwyer, Hogan and Stewart has said about the ease with which we read arguments as narratives (Dwyer, Hogan & Stewart, 2012). The example in figure 5.4 shows the incorrect narrativizing leads to a premise (It is bad for the environment) leads to a sub-conclusion (It is expensive), that is not the intended form of the argument, nor in any way a proper sub-conclusion. The two together however, leads to the sub-conclusion (I drive to school, but it isn't good) in the argument, and while it may not be a clear sub-conclusion given the premises, it does represent the argument stated.

(a) Correct interpretation

(b) Narrativizing interpretation

Figure 5.4: Narrativizing error

5.3 Perceived Difficulty

The students also filled out a custom questionnaire with four questions dealing with the perceived difficulty of different aspects of argument mapping. It was filled out by 20 out of 20 of the 8th graders and 24 out of 25 of the 7th graders. The questions were answered on a 5 point likert scale, ranging from very easy to very hard, as well as two long text answers

where they were given more room to explain what was hard and where they got stuck. The likert scale questions have been divided into Easy(1-2), Intermediate (3), and Hard (4-5).

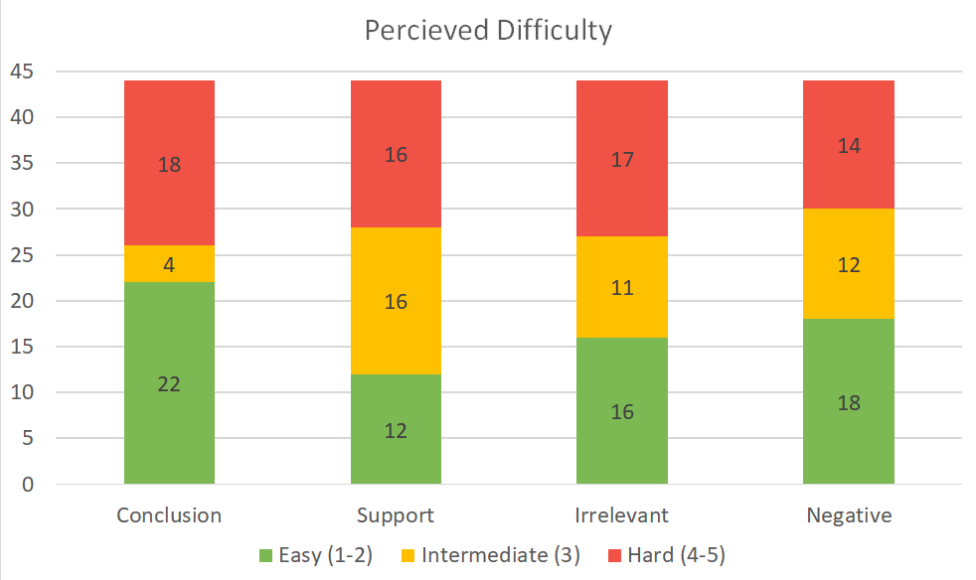

Figure 5.5: Perceived difficulty with argument mapping

As seen in figure 5.5, 41% (18/44) found it hard to identify the conclusion, while the data shows that the students were very adept at doing this, with only 23% either completely wrong or had an incorrect conclusion over the entire set of tasks. Even at the complex tasks, where the students had problems finding the correct structure, they still managed to identify the conclusion reliably. On the other hand, 41% (18/44) found it easy to find the negative premises, which was not supported by the data, where negative premises were a common error to make, at about 30% of the errors made.

Two long text answer questions were asked in addition to the four Likert scale questions: "Was there anything you got stuck on, or didn't understand?" and "Did you at any point feel like you needed more information about argument mapping, or more help in order to finish a task?". A quick overview of the answers were done in order to identify themes or recurring complaints. A majority of the answers were a short "no" or "yes", but a few of the answers showed two recurring themes: Difficulty in remembering the concepts, and difficulty with too many premises in the end.

Answers on the theme of "Difficulty in remembering the concepts":

- Swe: "ja, det var svårt att komma ihåg tex vad irrelevanta premisser var, alltså vad det innebär och vad det betyder."
  Eng: "yes, it was hard to remember what for example irrelevant premises were, like what it stands for and what it means."

- Swe: "Ja, man glömde lätt bort vad t.ex. negativa premisser hade för mening och begreppets betydelse"
  Eng: "Yes, you easily forgot what for example negative premises meant and the concepts meaning"

- Swe: "det kändes som om jag ibland glömde bort vad jag gjorde så jag skulle enkelt kunna gå tillbaka till instruktionerna."
  Eng: "it felt as if I sometimes forgot what I was doing and so I could easily go back to the instruction"

Answers on the theme of "Difficulty with too many premises in the end":

- Swe: "ja ibland kan jag tycka det. Då det i slutet var så många mer och det var svårt att urskilja vad som var vad ex. Slutsatsen."
  Eng: "yeah sometimes I can feel that. When in the end there were so many more and it was hard to separate what for example was the conclusion."

- Swe: "jag tyckte det gick väldigt bra i början men på slutet blev uppgifterna lite försvåra"
  Eng: "I thought it went very well in the beginning but towards the end the tasks were a bit too hard"

- Swe: "tyckte att del 4 [komplexa argument] blev betydligt svårare än dom andra"
  Eng: "I thought part 4 [the complex arguments] became significantly harder than the others"

6 | Discussion

The following goals of the thesis was stated in the introduction:

1. Design and evaluate a digital argument mapping tool in terms of usability
2. Identify if students are able to understand and perform basic argument mapping, and look at what kinds of errors they make

This chapter will first discuss the tool, followed by the argument mapping proficiency shown by the students, and finally the conclusions that can be drawn given the results of this thesis.

6.1 The Tool

The SUS score of 63.5, which lands in the "OK" to "GOOD" category, allows for an interpretation of the results that can not be dismissed due to a poorly working ealy version of the tool. The tool was however in a quite early stage, and lacked some clear feedback and scaffolding, which naturally results in a less comprehensive understanding of AM for the students, and also a lower SUS score due to inconsistencies in the tool. This coupled with the short time the students had to interact with the tool, and to learn about AM, slight caution should be exercised when interpreting the results, especially when trying to extrapolate too much from the results.

In the SUS questionnaire, question 6 (There is too much inconsistency in the tool) was given a much lower score than the other questions. When conducting the test several students asked what "inconsistent" meant, and said that they did not understand what it meant for the tool. Subsequently many scored a 3, which can be interpreted as a "safe" choice when you do not know what a question means. This may have given a lower score, based not on inconsistency in the tool, but on the students not understanding the question and choosing 3 as a default, but this hypothesis can not be further supported.

The final question "I need to learn much before I can become productive with the tool" also received a low score. This may be because the tool itself is made to be excessively difficult in the final tasks, in order to see what the students are having problems understanding. Therefore the score itself is not to be questioned, the students seemed to accurately identify the lack of training they had, which was as a result of the testing conditions.

One possible feature that was not implemented in the tool was user sessions, both in order to look at the tasks sequentially tied to one user, as well as connecting the results of the AM to the answers on the questionnaire. This was not implemented due to time constraints, as the testing sessions were moved to an earlier date than was originally planned for. User sessions would have allowed for more in depth analysis of individuals performance and connect it to their perceived difficulty. As the prototype stands, no connection can be made between tasks, questionnaire and users, nor can the tasks themselves be tied to one user.

The tool was in general useful to the students, up to a point, and then the need for feedback and more education in AM became clear.

6.1.1 Simplifications of argument mapping

The decision to simplify the tool was taken both to limit the evaluation to AM, in order not to require too much instruction in CT to test the structural aspects of AM, and to limit the amount of time spent developing the tool, as the tool had to be fully developed before being tested with the students. This resulted in a tool that skipped a few steps of the natural process of AM, specifically from text argument to premises, as well as a tool that did not do any of the evaluating of the AM when done. This leads to a conclusion about the efficacy of AM in secondary school that can only be based on the proficiency the students had for structuring AM, and not how well they actually will learn CT with an AM approach.

The tasks were modified after the pilot test, that was carried out with an adult and one appropriately aged individual, but the tasks did not receive any extensive validation before being used in the school. This could lead to some of the tasks having been too difficult, or having unclear language. In the evaluation of the finished argument maps, some interpretations that was not intended when writing the tasks were found, but these were not considered large deviations that would completely change the meanings of the arguments.

6.1.2 Future development of the tool

The specific features of the thesis tool that was identified as lacking, or features that should be focused on given further development of the tool (or possibly other tools with a similar intention):

- More comprehensive introductions and tutorials

Here the tool was lacking, as evident by the disconnect in the perceived difficulty, as well as the lower scores in complex arguments. Furthermore, a need exists to clarify not only the how of argument mapping, but the why as well. If the students do not understand why they are supposed to construct an argument map as close to the stated argument as possible, they may take shortcuts that fundamentally undermine the AM process.

- Focus on feedback and scaffolding within the tool

The prototype followed van Gelder's (2001) Quality Practice criteria, although clearly lacked somewhat in the feedback and scaffolding department. This is also supported by what Hattie & Timperley (2007) argued, that a much greater focus needs to be put on feedback for the tool to be reliably used to teach CT through AM.

- More training tasks in the tool, and more time to practice

Following the LAMP approach (van Gelder, 2001), more training in basic argument mapping, and dealing with more complex arguments should be implemented. This, coupled with increased interaction with the tool would hopefully result in increased proficiency in AM, as van Gelder claims.

- Automated feedback system for the tasks

In order to more easily deal with an increase in tasks, and make the possibility of integration into schools easier, an automated feedback system should be looked at, at least for the simpler tasks. The attempt Butchart et. al. (2009) conducted could be used as an inspiration, as such a system would be invaluable to teachers. What must not be forgotten is the quality of the feedback offered through such a system, as discussed earlier.

- A look at integration with existing courses

If attempting to have the tool actually used in schools, a hard look needs to be taken at in what capacity it can be used by the schools. If there is no larger project in motion where one can assume an entire new course will be given, there has to be some kind of integration with the existing curriculum. Candidate courses in secondary school would likely be History or Social Science, as they already discuss source criticism to some extent. Integration will have some bearing on both the content and the form of the tool, but would need to be further investigated depending on the context.

- Implement the full AM process, giving instruction and training, from natural language to the evaluation of argument maps

In current form, the prototype only deals with the mapping part of argument mapping. In order to give the full picture, there needs to be more connection to how AM is used from start to end. Here it is likely that some simplification still needs to be done, as for example the process of turning natural text into premises is not always easy, and many unstated premises can exist. However, when it comes to the evaluation of premises and arguments, more can be done in a tool to allow for the students to learn not only to construct maps, but to actually see if they are sound or not.

6.2 Argument Mapping and Task Difficulty

There was some difference between the two classes, that bore out on the results to some extent. The 8th grade class had a much louder volume when working, with students regularly talking and bothering each other. In general they needed a more hands on approach, with constant supervision. The 7th grade were mostly quiet throughout, and required very little supervision. The sample is quite small at 45 students, but still shows that not only calm and orderly classes can deal with these kinds of methods, but the less calm and less focused classes still are able to deal with the method adequately.

Looking at the most common problems when it comes to the complex tasks, both the irrelevant and the negative premises were difficult for the students. With the negative premises the students often missed to correctly identify the correct negative premise, and equally often had problems connecting the negative premise to the correct premise. It is possible that students tried to "solve" arguments, instead of trying to represent the stated argument, which may result in more sound arguments, but missing the point of constructing an AM in order to evaluate the structure of the stated argument. A possible, and even probable, cause of this is the lack of feedback, along with the general difficulty with interpreting an argument. As mentioned previously, this would likely need more AM training to overcome, as the LAMP approach recommends (van Gelder, 2001), and there is no reason to believe that with further training a higher level of mastery will be attained.

The tool itself makes only the briefest of introductions to AM (as seen in figures 3.3 and 3.4), and allows only for a few training tasks before starting the complex tasks. The result that the complex tasks would not to a high degree be correct was not surprising, but there is slightly more spread in the errors than was initially expected. This is also possibly due to the quick introduction, not giving the time for the strategies in tackling the larger arguments to be comprehensively taught in such a short time, as the students having only two scheduled 50 minute meetings with the classes may be too short a time to learn. In addition to this, the students had never had any formal education in critical thinking, logic, nor argument mapping, forcing the tool to bear the brunt of the teaching even in its early stage of development.

However, given the solid result in the basic tasks, as well as the consistently good identification of conclusions, the use of AM in secondary school should be ripe for further study. The students can do the basics of argument mapping, and with either a more developed tool, or more likely, with additional instruction in argument mapping and critical thinking in general, the students will be able to further gain understanding of the process. There is a possibility that the negative premises give an additional layer of complexity that the students would need additional time to properly understand. The complexity is due to not only having to identify the negative premise, but also identify to what premise it relates. Given more training, there is the possibility that this may lead to a better understanding of the concept, and improve their results.

With the relative size of the complex tasks that the students would engage in, as compared to the much more difficult tasks that university students deal with, it is also possible to design an automated feedback system, like the one Butchart et. al. attempted (2009). This would lessen the pressure and required competence by teachers in AM and CT, while freeing up time for discussion around the important pillars of CT in class. It would likely still be recommended to have some kind of more structured discussion outside of the tool regardlessly, as CT if it is to be transfered or generalized at all, needs to be discussed in a variety of contexts in addition to the training and practice delivered by AM.

To summarize, given that students can learn basic AM with very little instruction, and that studies have repeatedly shown that CT education through AM is effective, there should be no reason to assume that AM should not be further studied in secondary school.

6.2.1 Perceived difficulty with argument mapping

The reason to look at perceived difficulty was in order to find out if there was any disconnect between actual performance and perceived performance. If there is a large disconnect, there is likely some issues with feedback being too sparse, meaning that the students do not know if they are correct or not.

With the long text answers, the themes of “difficulty in remembering the concepts”, and “difficulty with too many premises in the end” were identified. Here it was clear that some students realized that the complex arguments were a bit too hard, which was intended. As these complex tasks were meant to push the limits of the student in order to identify where they lacked structure, or where they made mistakes, this theme in the answers was to be expected. The theme of “difficulty with remembering the concepts” however, is likely from the fact that introductions to these concepts was short and quick, and no supplemental information was given. This could however mean that when doing the complex tasks, the students could have potentially forgotten one concept, and therefore systematically missed all the possible instances of for example the negative premise.

Then the importance of the intended context becomes more clear. If a tool like this is intended to be used completely isolated, without other courses or information about AM or CT, there

would be much more need for feedback and scaffolding. However, if it is possible to use the tool as a complement to other general teaching of CT, and perhaps some AM, the need for feedback and scaffolding lessens to some degree. In any case, the importance of proper feedback within the tool can not be ignored, as Hattie & Timperley point out in their seminal paper The Power of Feedback (Hattie & Timperley, 2007).

6.2.2 Future study of argument mapping in secondary school

In addition, a few general recommendations for studies concerning AM and CT in secondary school has been noted:

- Running longer CT studies using a digital AM tool for secondary school students

In order to allow more time for training, no matter what study is conducted, one should allow for more training, adhering to the LAMP approach (van Gelder, 2001) as mentioned earlier. There is also a non-trivial task of determining what the best use of a digital AM tool would be, from simple complement in class to completely freestanding tool, which would in itself affect the amount of training and instruction in a tool. This would have to be taken into consideration when designing an experiment, or when designing a tool to be tested.

- Study the appropriate complexity of arguments for students in secondary school

When trying to teach CT through AM, some simplification needs to be done for secondary school, but exactly what these simplifications would be needs to be studied further. For example a slower scaling in complex tasks, tested on a larger set of students could give a baseline of what the students find graspable, both in terms of length of arguments, and complexity of the arguments themselves.

- Studies attempting to measure actual CT increase with the use of digital AM tools in secondary school

This follows the general format of studies conducted at university levels, and given the results of this thesis, there seems to be no reason to doubt AM's effectiveness, therefore studies need to look at to what extent the effect is replicated in younger ages. This may still need a custom digital AM tool to do, but it could be possible with standard tool such as Rationale (van Gelder, 2002), if the accompanying course takes the complexity of the tool into consideration. As mentioned in the thesis, tests such as the HCTA (Halpern, 2018) or the CCTST (Facione, 2018), were deemed somewhat too complex to be able to yield any useful results in this thesis, but if the students have more time, it is possible that these tests could be used, or possibly some kind of simplified versions of these tests.

Another possibility is to use tests designed to test similar features as the CT tests, but simpler, or made for younger subjects. An example of this is the Lectical Reflective Judgment Assessment (Dawson, 2008), a test that has been shown to have some correlation with CT, and to CT disposition (Dwyer, Hogan & Stewart, 2015). Another example may be the Motivated Strategies for Learning Questionnaire (Pintrich & DeGroot, 1990), that measures motivation, elaboration and self-regulation, which are all parts of the disposition to CT. With tests such as these, you are not testing CT directly, but it may prove more fruitful with younger students than the language heavy CT tests.

6.3 Conclusion

The tool is a functioning but still early first foray into a simplified argument mapping software for secondary school, given the SUS evaluation and the level at which the students managed to do argument mapping with it. However, the implementation of a simplified tool is only a first dip of the toe into the waters of critical thinking education in secondary school. The general conclusions to draw from this explorative study supports that students in secondary school can use argument mapping in a basic way, and can do so after less than two hours of training and limited instruction.

Given that there is support for AM being used to teach CT in general (Dwyer, Hogan & Stewart, 2010, 2012, 2015; van Gelder, 2015; Twardy, 2004; Alvarez Ortiz, 2007), and a drive for more critical thinking in schools (Swedish National Agency for Education, 2011), as well as the fact that the thesis has shown a potential for secondary school students to use and manipulate arguments with argument mapping, further study of this should be considered promising. As our young meet more and more complex information, and are told not to trust everything they hear, they need to be given this groundwork for a critical examination of information as soon as possible.

In conclusion, there is clear potential in the use of AM in CT education in secondary school, and further studies should be conducted in order to help develop a tool or a course that can be used by students to get a head start on their current and future needs for critical thinking in our new, information dense world.

Bibliography

Alvarez Ortiz, M. (2007). "Does philosophy improve critical thinking skills?" PhD thesis.

Bangor, A., J. Miller & P. Kortum (2009). "Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale". In: *Journal of Usability Studies* 4, p. 3.

Butchart, S., D. Forster, I. Gold, J. Bigelow, K. Korb, G. Oppy, & A. Serrenti (2009). "Improving critical thinking using web based argument mapping exercises with automated feedback". In: *Australasian Journal of Educational Technology* 25, p. 2.

Cavanagh, R. (2015). *Argument Mapping with Rationale*. URL: https://oddrops.blog/2015/09/24/argument-mapping-with-rationale/.

Dawson, T. L. (2008). *The Lectical Assessment System: Validity and reliability*. Northhampton, MA: Developmental Testing Service, LLC.

Dwyer, C. P., M. J. Hogan & I. Stewart (2010). "The evaluation of argument mapping as a learning tool: Comparing the effects of map reading versus text reading on comprehension and recall of arguments". In: *Thinking Skills and Creativity* 5, pp. 16–22.

— (2012). "An evaluation of argument mapping as a method of enhancing critical thinking performance in e-learning environments". In: *Metacognition and Learning* 7.3, pp. 219–244.

— (2015). "The effects of argument mapping-infused critical thinking instruction on reflective judgement performance". In: *Thinking Skills and Creativity* 16, pp. 11–26.

Ennis, R. H. (1987). "A taxonomy of critical thinking dispositions and abilities". In: *Teaching thinking skills: Theory and practice*. Ed. by J. B. Baron & R. J. Sternberg. New York: Times Books, pp. 9–26.

Facione, P. (1989). "Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction". In: *Research Findings and Recommendations* 315.

— (2018). *California Critical Thinking Skills Test (CCTST)*. URL: https://www.insightassessment.com/Products/Products-Summary/Critical-Thinking-Skills-Tests/California-Critical-Thinking-Skills-Test-CCTST.

Florea, N. M. & E. Hurjui (2015). "Critical Thinking in Elementary School Children". In: *Procedia - Social and Behavioral Sciences* 180, pp. 565–572.

Halpern, D. F. (1998). "Teaching critical thinking for transfer across domains". In: *American Psychologist* 53.4, pp. 449–455.

— (2018). *Halpern Critical Thinking Assessment*. URL: https://sites.google.com/site/dianehalperncmc//home/research/halpern-critical-thinking-assessment.

Hattie, J. & H. Timperley (2007). "The Power of Feedback". In: *Review of Educational Research* 77.1, pp. 81–112.

Holmes, N. G., Carl E. Wieman & D. A. Bonn (2015). "Teaching critical thinking". In: *Proceedings of the National Academy of Sciences* 112.36, pp. 11199–11204.

Inhelder, B. & J. Piaget (1958). *The growth of logical thinking: From childhood to adolescence*. New York: NY, US: Basic Books.

Korb, K. B. (2014). *The Process of Argument Analysis*. URL: https://bayesianwatch.wordpress.com/2014/04/18/tpaa/.

Lipman, M. (1976). "Philosophy For Children". In: *Metaphilosophy* 7.1, pp. 17–33.

Lipman, M. (1988). "Critical Thinking and the Use of Criteria". In: *Inquiry : An Interdisciplinary Journal of Philosophy* 1.2, pp. 2–2.

Liu, O. L., L. Frankel & C. Roohr (2014). "Assessing Critical Thinking in Higher Education: Current State and Directions for Next-Generation Assessment". In: *ETS Research Report Series* 2014.1, pp. 1–23.

Northwoods Software (2018). *GoJS*. URL: https://gojs.net/.

Paul, Richard (1992). "Critical thinking: What, why, and how". In: *New Directions for Community Colleges* 1992.77, pp. 3–24.

Pintrich, R. R. & E. V. DeGroot (1990). "Motivational and self-regulated learning components of classroom academic performance". In: *Journal of Educational Psychology* 82, pp. 33–40.

Reasoning Lab (2018). *Bcisive - online decision mapping*. URL: https://www.bcisiveonline.com/.

Schneider, D. C., C. Voigt & G. Betz (2007). "ArguNet-a software tool for collaborative argumentation analysis and research". In: *7th Workshop on Computational Models of Natural Argument (CMNA VII)*.

Scriven, Michael (1979). *Reasoning*. McGraw-Hill.

Swedish National Agency for Education (2011). *Skolans värdegrund och uppdrag*. URL: https://www.skolverket.se/laroplaner-amnen-och-kurser/grundskoleutbildning/grundskola/laroplan/curriculmtopdf.pdf?typeOfSchooling=gr.

Twardy, C. (2004). "Argument Maps Improve Critical Thinking". In: *Teaching Philosophy* 27.2, pp. 95–116.

van Gelder, T. J. (2001). "How to improve critical thinking using educational technology". In: *Meeting at the Crossroads. Proceedings of the 18th Annual Conference of the Australian Society for Computers In Learning In Tertiary Education (ASCILITE 2001)*. Ed. by M. Keppell Kennedy, C. McNaught & T. Petrovic. Biomedical Multimedia Unit, The University of Melbourne, pp. 539–548.

— (2002). *Argument mapping with Reason!Able*. American Philosophical Association Newsletter on Philosophy and Computers.

— (2015). *Using Argument Mapping to Improve Critical Thinking Skills*. The Palgrave Handbook of Critical Thinking in Higher Education.

Whately, R. (1831). *Elements of logic*. B. Fellowes.

A | Appendix A

A.1 Data Table

Task	Compl Correct	Compl Wrong	Corr. Concl.	Wrong Concl.	Corr. Irr.	Wrong Irr.	Corr. Neg.	Wrong Neg.	Wrong Sup.	Incomp	Total
Basic1	42	2	0	0	0	0	0	0	0	1	45
Basic2	41	3	0	0	0	0	0	0	0	1	45
Irr1	35	1	0	7	4	2	0	0	0	2	45
Irr2	36	2	3	1	0	4	0	0	0	3	45
Neg1	24	6	10	5	0	1	4	10	2	0	45
Neg2	17	8	16	1	0	0	0	17	3	3	45
Compl1	2	7	35	0	0	21	6	25	13	1	45
Compl2	11	8	24	1	10	18	9	15	21	1	45
Compl3	8	10	24	2	4	6	12	12	21	1	45
Compl4	3	4	34	1	2	14	17	15	25	3	45
Compl5	7	9	24	0	2	10	2	5	18	4	44
Compl6	6	14	16	2	0	12	10	8	20	5	43
Compl7	1	7	30	0	1	7	9	21	27	5	43
Compl8	2	9	24	0	2	22	7	17	22	4	39
Total	235	90	240	20	25	117	76	145	172	34	619

Figure A.1: Task Data

B | Appendix B

B.1 Tasks: Basic

Figure B.1

Alla stolar har fyra ben. Jag har en stol. Därför har min stol har fyra ben.

All chairs have four legs. I have a chair. Therefore, my chair has four legs.

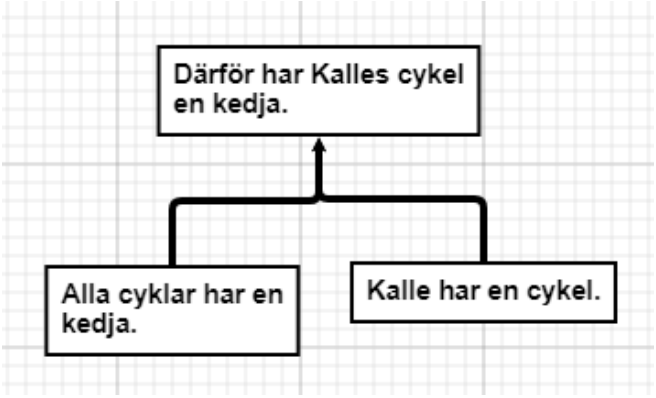

Figure B.2

Alla cyklar har en kedja. Kalle har en cykel. Därför har Kalles cykel en kedja.

All bikes have a chain. Kalle has a bike. Therefore, Kalles bike has a chain.

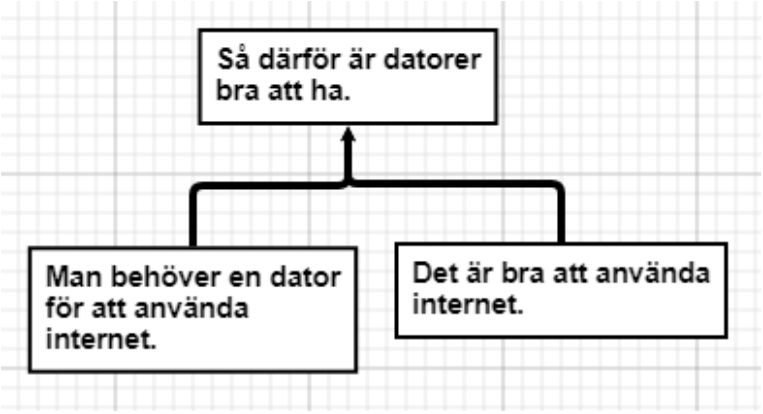

Figure B.3

Det är bra att använda internet. Man behöver en dator för att använda internet. Så därför är datorer bra att ha.

It's good to use the internet. You need a computer to use the internet. Therefore computers are good to have.

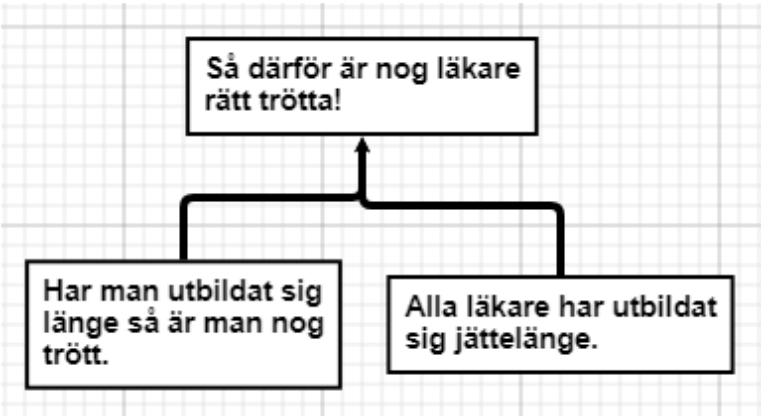

Figure B.4

Alla läkare har utbildat sig jättelänge. Har man utbildat sig länge så är man nog trött. Så därför är nog läkare rätt trötta!

All doctors have studied for a long time. If you've studied for a long time you're probably tired. So that's probably why doctors are tired.

Figure B.5

Får och getter är rätt lika. Saker som är lika äter oftast samma saker. Då tror jag att får och getter äter samma saker.

Sheep and goats are pretty similar. Things that are similar usually eat the same things. So I think sheep and goats eat the same things.

B.2 Tasks: Irrelevant

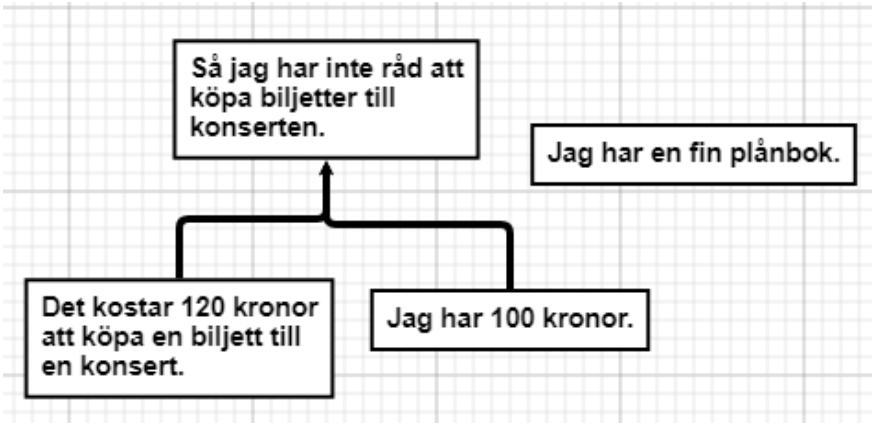

Figure B.6

Jag har 100 kronor. Det kostar 120 kronor att köpa en biljett till en konsert. Jag har en fin plånbok. Så jag har inte råd att köpa biljetter till konserten.

I have 100 kronors. It's 120 kronors to buy a ticket to a concert I have a nice wallet. So I can't afford to buy a ticket to the concert.

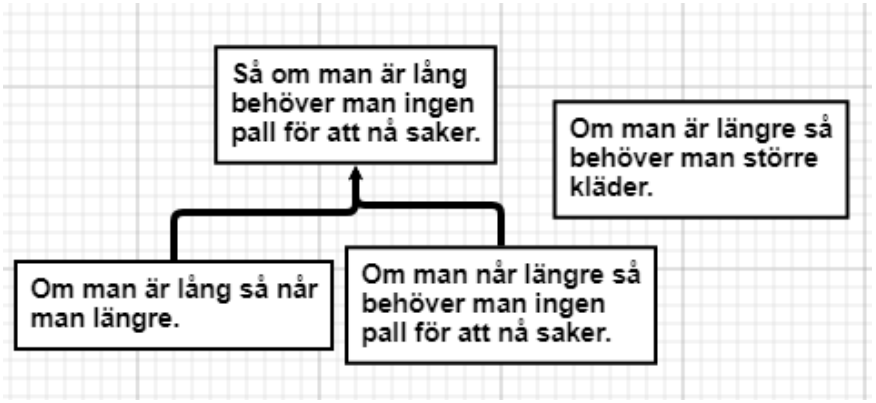

Figure B.7

Om man är lång så når man längre. Om man når längre så behöver man ingen pall för att nå saker. Om man är längre så behöver man större kläder. Så om man är lång behöver man ingen pall för att nå saker.

If you're tall you can reach further. Of you can reach further you don't need a chair to reach things. If you're tall you don't need clothes. So if you're tall you don't need a chair to reach things.

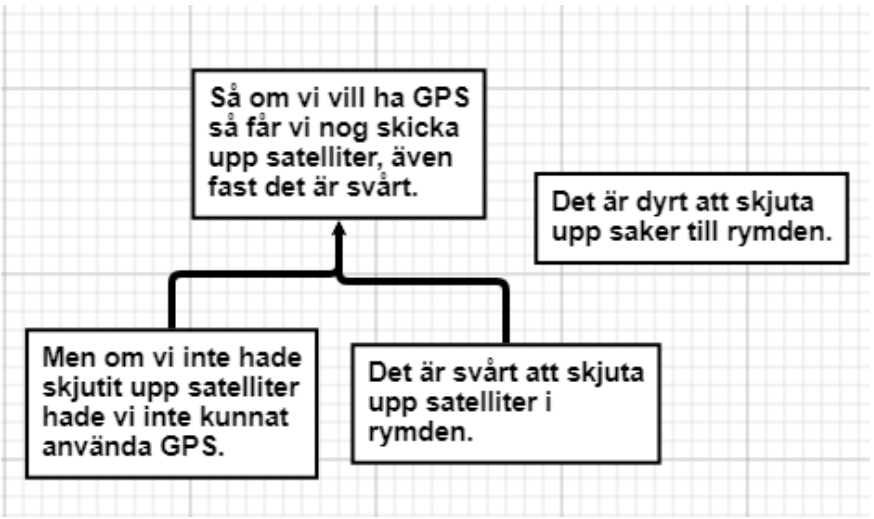

Figure B.8

Det är svårt att skjuta upp satelliter i rymden. Men om vi inte hade skjutit upp satelliter hade vi inte kunnat använda GPS. Det är dyrt att skjuta upp saker till rymden. Så om vi vill ha GPS så får vi nog skicka upp satelliter, även fast det är svårt.

It's hard to launch satellites into space. But if we didn't have satellites we wouldn't be able to use GPS. It's expensive to launch things into space. So if we want GPS, we probably have to launch satellites, even though it's hard.

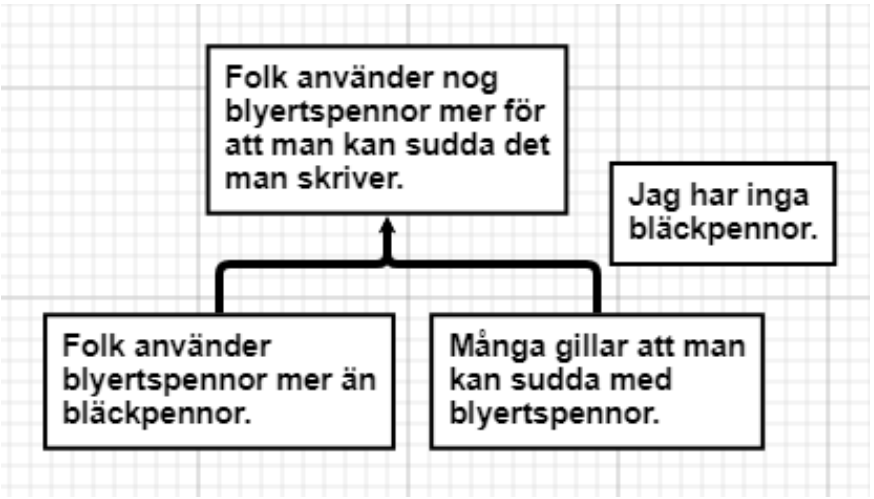

Figure B.9

Folk använder blyertspennor mer än bläckpennor. Många gillar att man kan sudda med blyertspennor. Jag har inga bläckpennor. Så jag tror att folk gillar blyertspennor för att man kan sudda det man skriver.

People use pencils more than pens. A lot of people like that you can erase what you write with pencils. I don't have any pens. So I think people like pencils because you can erase what you write with them.

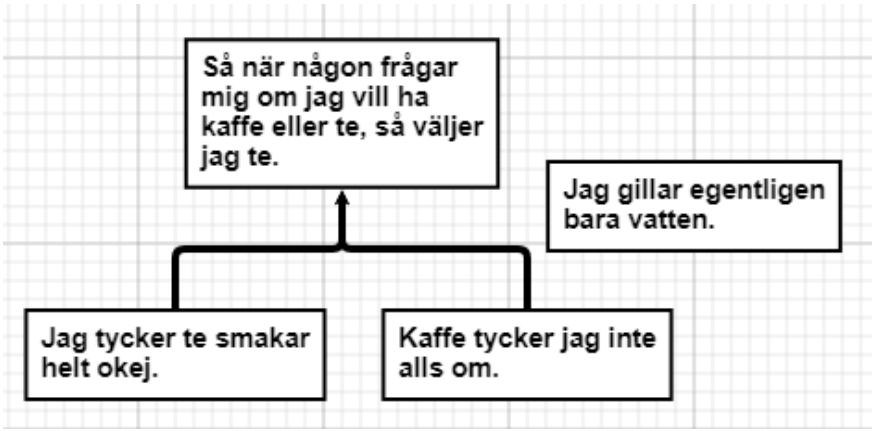

Figure B.10

Jag tycker te smakar helt okej. Kaffe tycker jag inte alls om. Jag gillar egentligen bara vatten. Så när någon frågar mig om jag vill ha kaffe eller te, så väljer jag te.

I think tea tastes pretty okay. I don't like coffee at all. I really only like water. So if someone asks me if I want coffee or tea, I'll chose tea.

B.3 Tasks: Negative

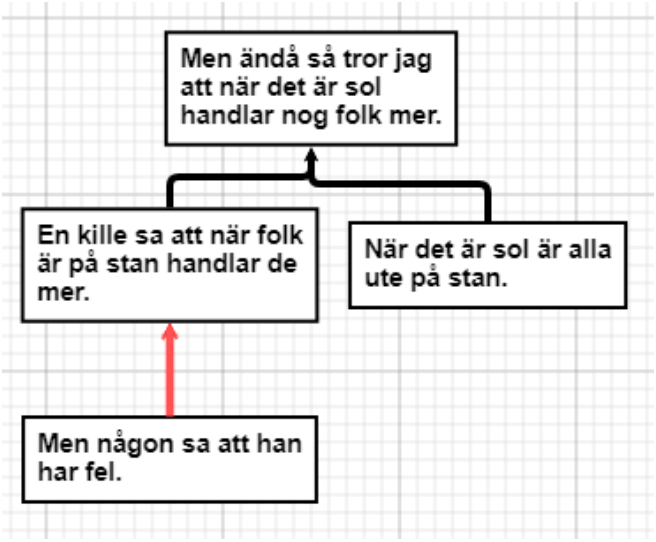

Figure B.11

När det är sol är alla ute på stan. En kille sa att när folk är på stan handlar de mer. Men jag tror att han har fel. Men ändå så tror jag att när det är sol handlar nog folk mer.

When it's sunny out everyone is out on the town. Some guy said that when people are out on the town they shop more. But I think he's wrong. But I still think that when it's sunny, people shop more.

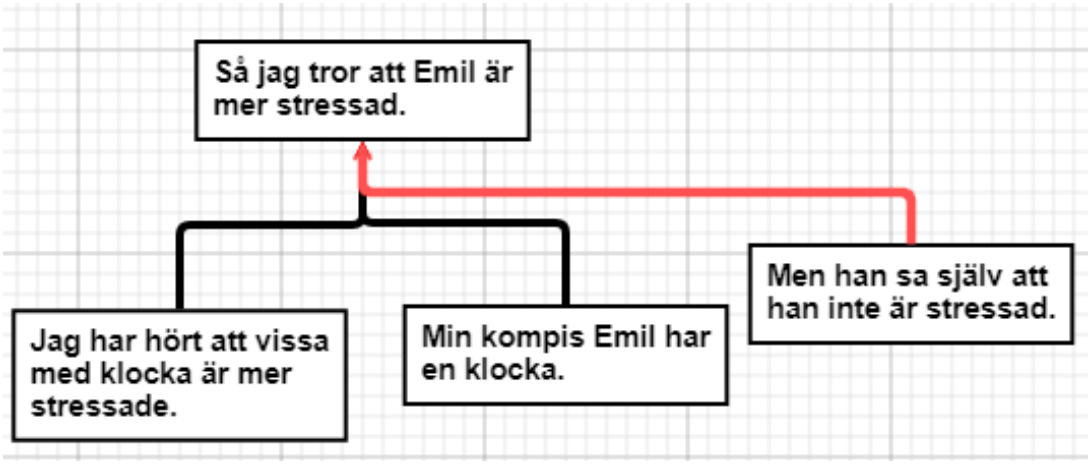

Figure B.12

Jag har hört att vissa med klocka är mer stressade. Min kompis Emil har en klocka. Så jag tror att Emil är mer stressad. Men han sa själv att han inte är stressad.

I've heard that some people who wear watches are more stressed. My friend Emil has a watch. So I think Emil is more stressed. But he said himself that he is not stressed.

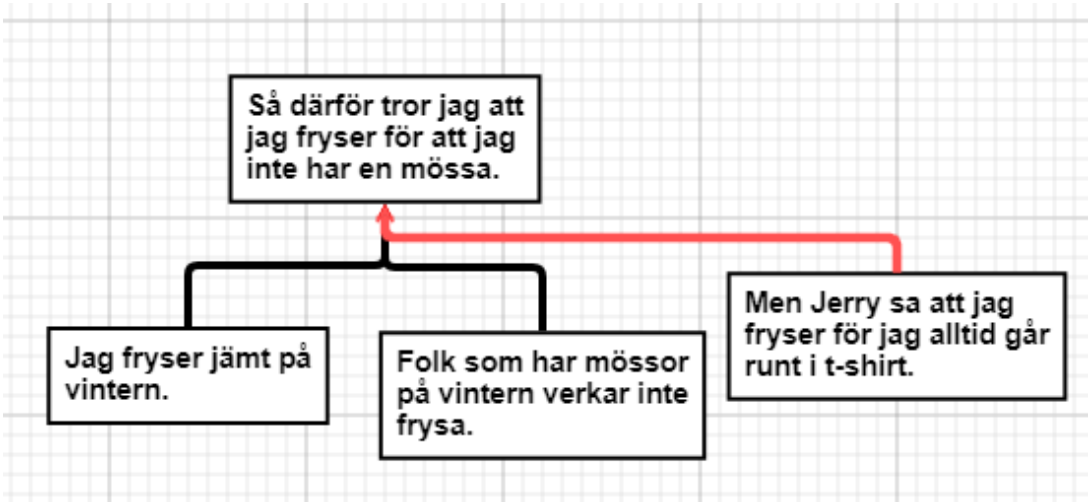

Figure B.13

Folk som har mössor på vintern verkar inte frysa. Jag fryser jämt på vintern. Så därför tror jag att det är nog för att jag inte har en mössa. Men Jerry sa att jag inte fryser för att jag inte har en mössa, jag fryser för jag alltid går runt i t-shirt.

People who wear hats during the winter do not seem to be cold. I'm always cold during the winter. That's probably because I don't have a hat. But Jerry said that I'm not cold because I don't wear a hat, I'm cold because I always wear a t-shirt.

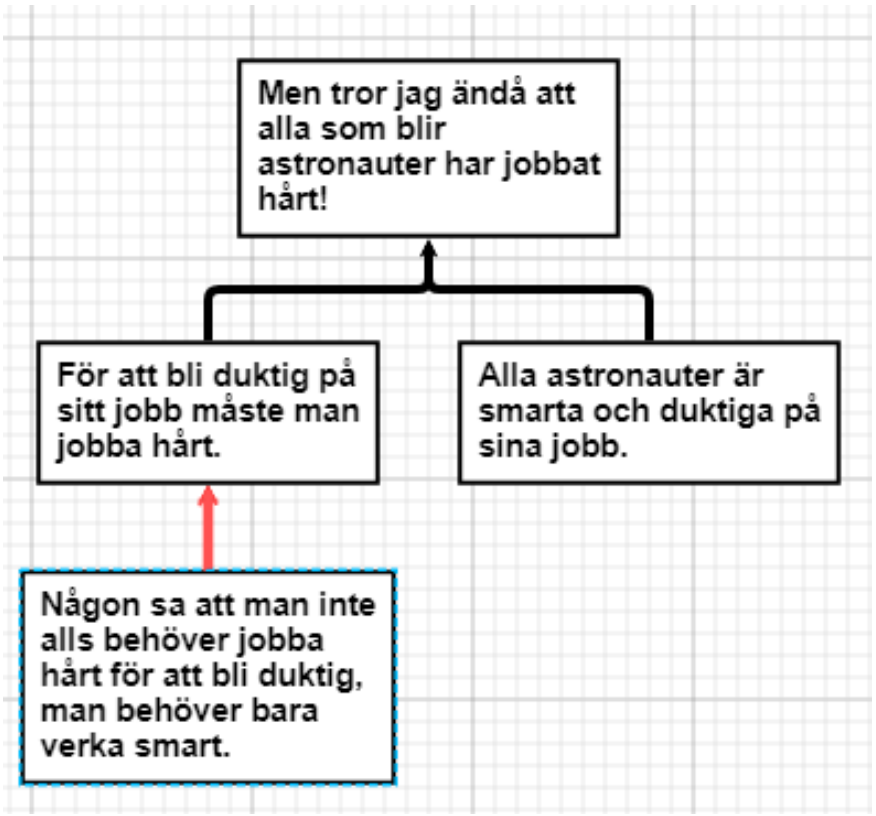

Figure B.14

Alla astronauter är smarta och duktiga på sina jobb. För att bli duktig på sitt jobb måste man jobba hårt. Någon sa att man inte alls behöver jobba hårt för att bli duktig, man behöver bara verka smart. Men tror jag ändå att alla som blir astronauter har jobbat hårt!

All astronauts are smart and good at their jobs. To become good at a job, you have to work hard. Someone said that you don't have to work hard to become good, you only need to seem smart. But I still think everyone who becomes an astronaut has worked hard!

Figure B.15

När jag sitter på bussen ser jag alltid barn med ryggsäckar. Jag hörde någon säga att alla barn med ryggsäckar är på väg till skolan. Men det kan ju inte vara rätt, för de kanske ska hem. Så därför tror jag inte att alla barn på bussen med ryggsäckar är på väg till skolan.

When I'm on the bus, I always see children with backpacks. I heard someone say that kids with backpacks are on their way to school. But that can't be right, they may be going home. So that's why I don't think every kid on a bus with a backpack is on their way to school.

B.4 Tasks: Complex

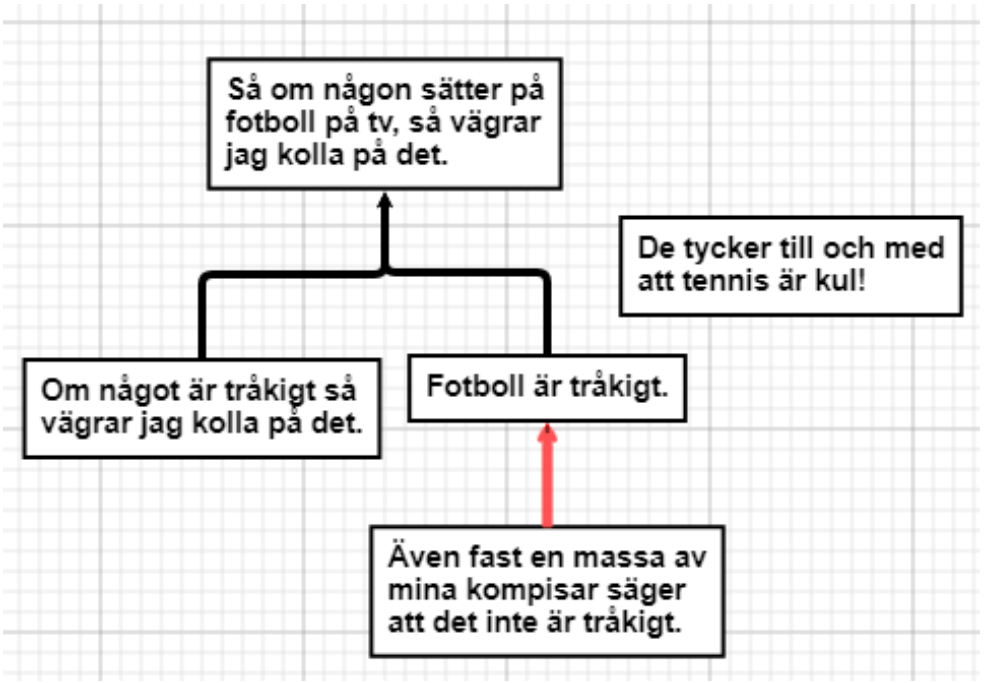

Figure B.16

Fotboll är tråkigt. Om något är tråkigt så vägrar jag kolla på det.Även fast en massa av mina kompisar säger att det inte är tråkigt. De tycker till och med att tennis är kul! Så om någon sätter på fotboll på TV, så vägrar jag kolla på det.

Soccer is boring. If something is boring I refuse to watch it. Even though my friends say it's not boring. They even find tennis fun! So if someone shows soccer on the TV, I refuse to watch it.

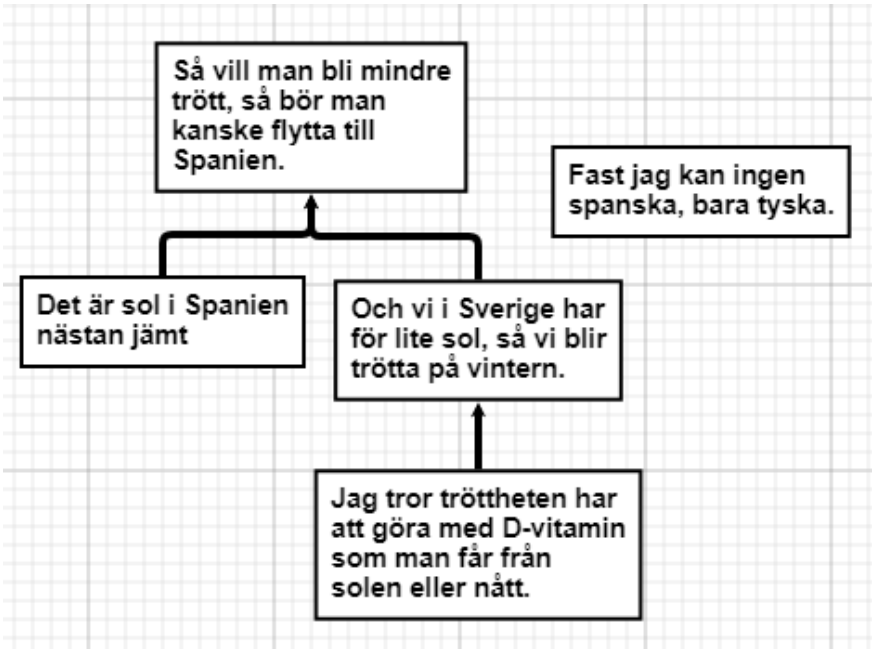

Figure B.17

Det är sol i Spanien nästan jämt. Och vi i Sverige har för lite sol, så vi blir trötta på vintern. Jag tror tröttheten har att göra med D-vitamin som man får från solen eller nått. Så vill man bli mindre trött, så bör man kanske flytta till Spanien. Fast jag kan ingen spanska, bara tyska.

It's sunny in Spain almost always. And in Sweden we don't have enough sun, so we get tired in the winter. I think the tiredness has to do with D-vitamin that you get from the sun or something. So if you want to be less tired, you should move to Spain. Although I can't speak Spanish, only German.

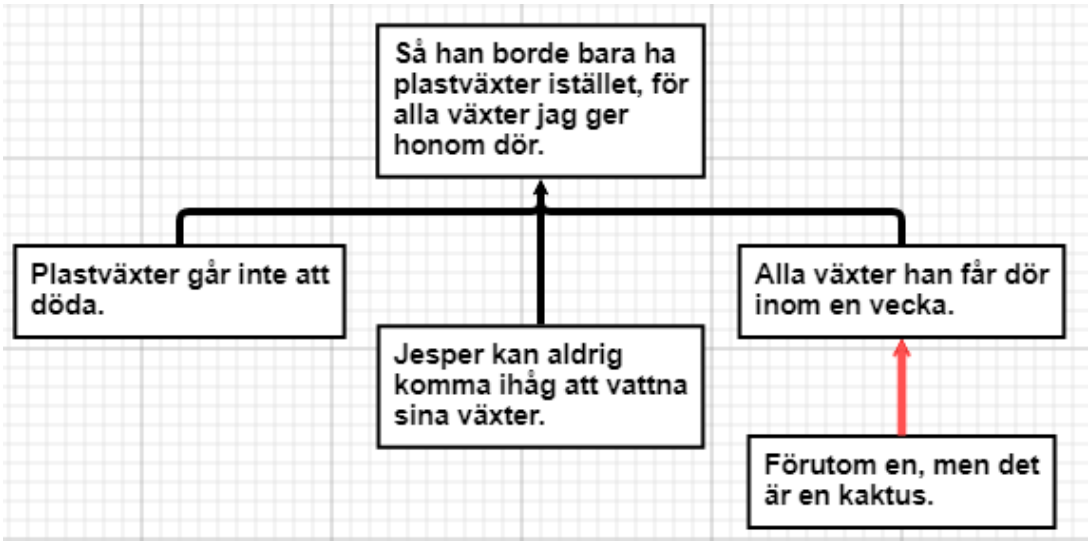

Figure B.18

Jesper kan aldrig komma ihåg att vattna sina växter. Alla växter han får dör inom en vecka. Förutom en, men det är en kaktus. Plastväxter går inte att döda. Så han borde bara ha plastväxter istället, för alla växter jag ger honom dör.

Jesper can never remember to water his plants. All plants he gets dies within a week. Except for one, but it's a cactus. Plastic plants can't be killed. So he should only have plastic plants instead, because all plants I give him die.

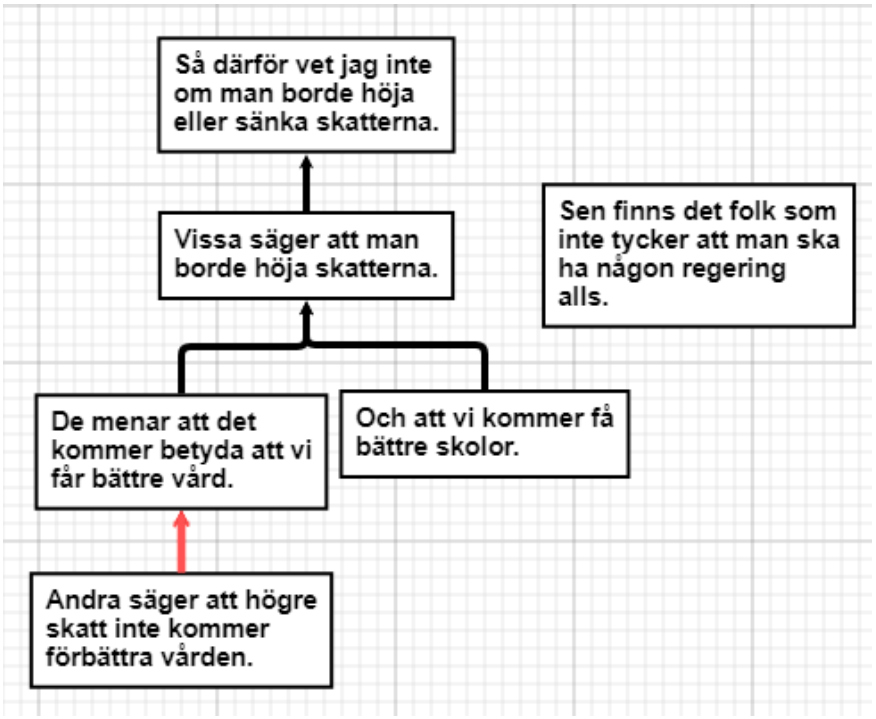

Figure B.19

Vissa säger att man borde höja skatterna. De menar att det kommer betyda att vi får bättre vård. Och att vi kommer få bättre skolor. Andra säger att högre skatt inte kommer förbättra vården. Sen finns det folk som inte tycker att man ska ha någon regering alls. Så därför vet jag inte om man borde höja eller sänka skatterna.

Some say we should increase the taxes. They say it will mean we get better health care. And that we'll get better schools. Others say that increased taxes won't improve health care. Then there's people that don't think we should have any government at all. So that's why I don't know if we should increase or lower taxes.

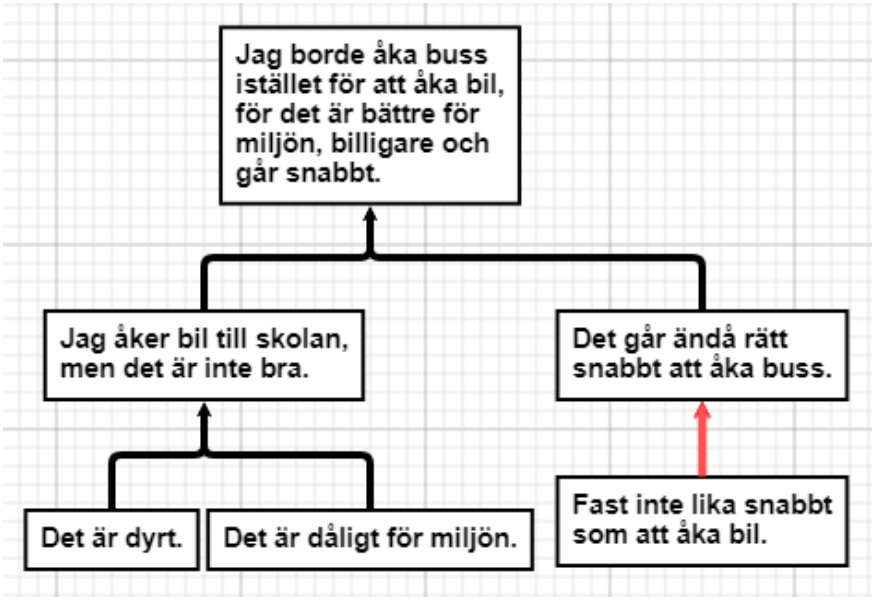

Figure B.20

Jag åker bil till skolan, men det är inte bra. Det är dyrt. Det är dåligt för miljön. Det går ändå rätt snabbt att åka buss. Fast inte lika snabbt som att åka bil. Jag borde åka buss istället för att åka bil, för det är bättre för miljön, billigare och går snabbt.

I take the car to school, but it's not good. It's expensive. It's bad for the environment. It's still pretty fast to ride the bus. Although not as fast as driving the car. I should ride the bus instead of driving the car, because it's better for the environment, cheaper, and is fast.

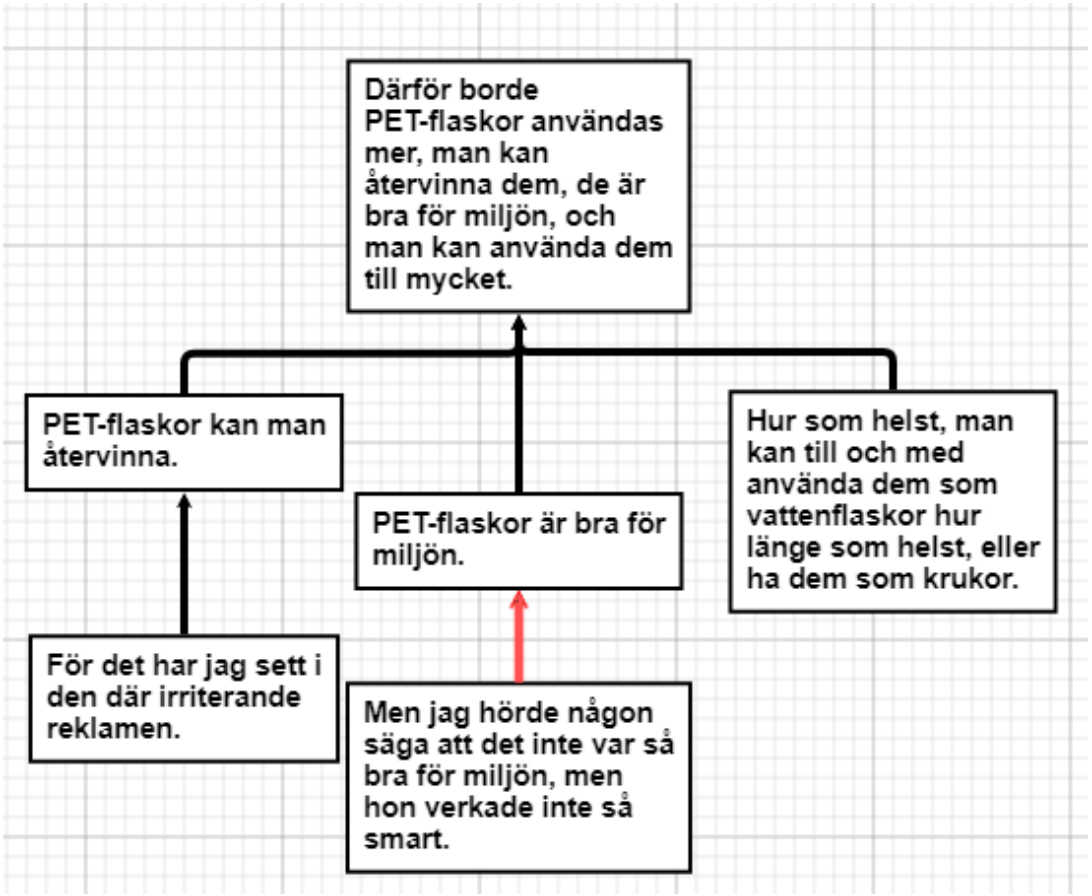

Figure B.21

PET-flaskor kan man återvinna. För det har jag sett i den där irriterande reklamen. PET-flaskor är bra för miljön. Men jag hörde någon säga att det inte var så bra för miljön, men hon verkade inte så smart. Hur som helst, man kan till och med använda dem som vattenflaskor hur länge som helst, eller ha dem som krukor. Därför borde PET-flaskor användas mer, man kan återvinna dem, de är bra för miljön, och man kan använda dem till mycket.

PET-bottles can be recycled. I've seen it in that annoying ad. PET-bottles are good for the environment. But I heard someone say that it wasn't so great or the environment, but she didn't seem that smart. Anyway, you can even use them as water bottles for a long time, or use them as pots. That's why PET-bottles should be used more, you can recycle them, they are good for the environment, and you can use them for a lot of stuff.

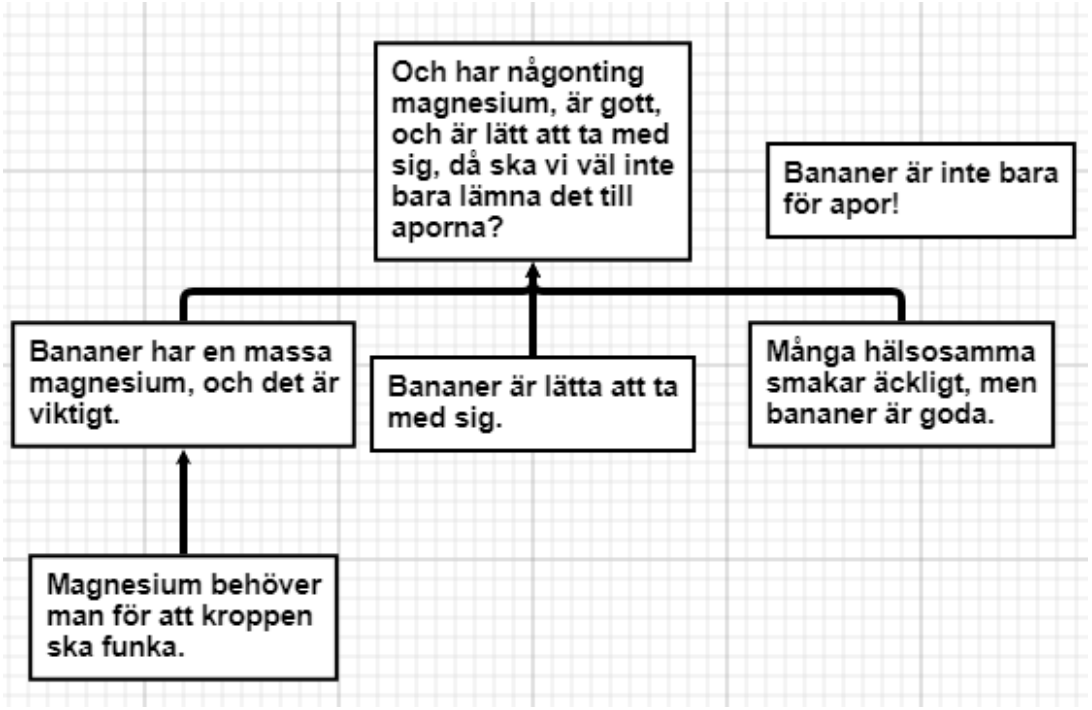

Figure B.22

Bananer är inte bara för apor! Bananer har en massa magnesium, och det är viktigt. Magnesium behöver man för att kroppen ska funka. Många hälsosamma saker smakar äckligt, men bananer är goda. Bananer är lätta att ta med sig. Och har någonting magnesium, är gott, och är lätt att ta med sig, då ska vi väl inte bara lämna det till aporna?

Bananas aren't just for monkeys! Bananas have loads of magnesium, and that's important. You need magnesium in order for the body to work. A lot of healthy things taste bad, but bananas are good. Bananas are easy to bring with you. And if something has magnesium, is good, and is easy to bring with you, then we shouldn't leave it to the monkeys, right?

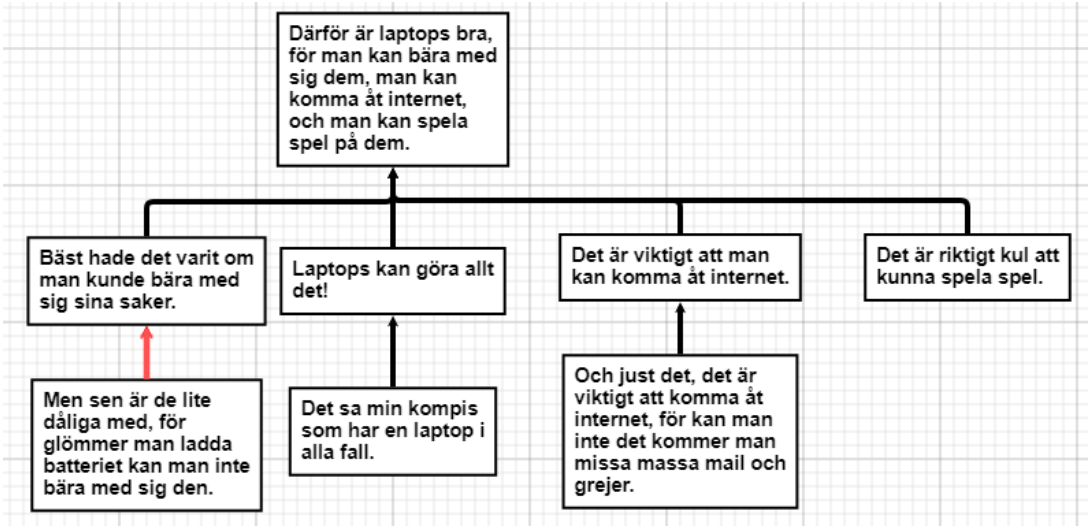

Figure B.23

Det är viktigt att man kan komma åt internet. Det är riktigt kul att kunna spela spel. Bäst hade det varit om man kunde bära med sig sina saker. Laptops kan göra allt det! Det sa min kompis som har en laptop i alla fall. Därför är laptops bra, för man kan bära med sig dem, man kan komma åt internet, och man kan spela spel på dem. Men sen är de lite dåliga med, för glömmer man ladda batteriet kan man inte bära med sig den. Och just det, det är viktigt att komma åt internet, för kan man inte det kommer man missa massa mail och grejer.

It's important to be able to have internet. It's really fun to play video games. It's great if you can carry things with you. Laptops can do all that! that's what my friend with a laptop said anyways. That's why laptops are great, you can carry them with you, you can get to the internet, and you can play games on them. But then they are a bit bad as well, because if you forget to charge the battery, you can't carry it with you. Oh, right, and it's important to get to the internet, because if you can't, you'll miss emails and stuff.

Figure B.24

Jag har sett en massa filmer om piloter. Har man sett en massa piloter som flyger så lär man sig säkert något. Jag har suttit i en cockpit en gång med, och sett alla spakar och grejer. Piloterna sa att det var rätt svårt att lära sig bara från att sitta i cockpit en gång. Men jag tror inte att det kan vara så farligt svårt. Så jag tror ändå att jag hade kunnat landa ett flygplan, jag har både sett en massa filmer och varit i en cockpit.

I've seen a bunch of movies about pilots. If you've seen a bunch of movies about pilots flying, you probably learn something. I've been in a cockpit once too, and seen all the buttons and stuff. The pilots said it's probably pretty hard to learn anything from just sitting in a cockpit once. But I don't think it can be that hard. So I still think I could land a plane, I've seen a bunch of movies and been in a cockpit.

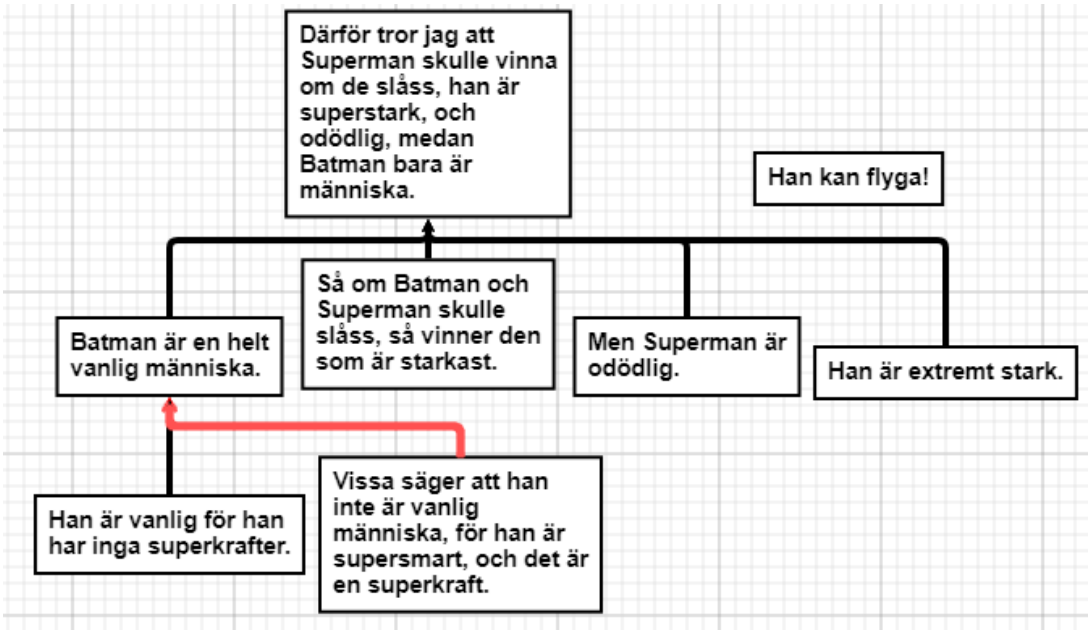

Figure B.25

Batman är en helt vanlig människa. Han är vanlig för han har inga superkrafter. Vissa säger att han inte är vanlig människa, för han är supersmart, och det är en superkraft. Men Superman är odödlig. Han är extremt stark. Han kan flyga! Så om Batman och Superman skulle slåss, så vinner den som är starkast. Därför tror jag att Superman skulle vinna om de slåss, han är superstark, och odödlig, medan Batman bara är människa.

Batman is a regular human. He's regular because he doesn't have any superpowers. Some say he's not a regular human, because he's super smart, and that's a superpower. But Superman is immortal. He's extremely strong. He can fly! So if Batman and Superman would fight, the strongest would win. That's why I think that Superman would win if they fight, he's extremely strong, and immortal, while Batman is just a regular human.