Linköping Studies in Science and Technology
Dissertation No. 2122

Crossing the threshold

Visualization design and conceptual
understanding of evolution

Andreas Göransson

Crossing the threshold

Visualization design and conceptual understanding of evolution

Andreas Göransson

Department of Science and Technology, Division of
Technical faculty
Linköpings universitet, SE-581 83 Linköping, Sweden
Linköping 2021

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

<https://creativecommons.org/licenses/by-nc/4.0/>

Cover art

Front: Mineral skeletons of radiolarians (Ernst Haeckel)

Back: Trilobite fossil drawing (Ernst Haeckel)

Ernst Haeckel (1834-1919) was an influential German zoologist. His works on radiolarians stands out, involving the naming of close to 150 new species. Haeckel was not only a zoologist but also a skillful artist and illustrator. The most famous of his publications is “*Kunstformen der Natur*” (“*Art Forms in Nature*”), including over one hundred prints of different species. The influence of the book was not limited to science but extended to contemporary architecture, art and design in line with Art Nouveau, for example the main portal of the Paris Exhibition 1900.

Radiolaria is a phylum of zooplankton, single-celled organism with a mineral skeleton often made of silica. It is the shape of these mineral skeletons that are depicted in Haeckel’s drawing on the front page.

Trilobites are an extinct class of marine arthropods evolved under the Cambrian period (540 – 485 mya).

© Andreas Göransson, 2021 (unless stated otherwise)

Printed in Linköping, Sweden by LiU-tryck, 2021

ISSN: 0345-7524

ISBN: 978-91-7929-707-7

Abstract

The theory of evolution is considered the unifying theory of biology. An accurate understanding of evolution is vital both for the understanding of diverse topics in biology, but also for societal issues such as antibiotic resistance or biodiversity. In contrast, decades of research in science education have revealed that students have difficulties to accurately understand evolutionary processes such as mutation and natural selection. The majority of this research relies on a conceptual framework of so-called key concepts (variation, selection, inheritance), derived from scholarly descriptions of natural selection. Recent research suggests that non-domain specific concepts such as randomness, probability, spatial and temporal scales, so called threshold concepts, are important for evolution understanding in addition to the key concepts. Thus, many important elements of evolutionary theory are counter-intuitive or lie outside direct perception. Hence, representations such as visualizations, models and simulations are considered to be important for teaching and learning evolution. While the importance of visualizations is generally acknowledged for science education, less is known about how visual design can facilitate students understanding of threshold concepts, such as random mutations or spatial scales.

This thesis uses the Model of Educational Reconstruction (MER) as the guiding framework for exploring the significance of threshold concepts by analysing the conceptual content of students' explanations and extant visualizations of natural selection. MER combines scientific content with teaching and learning perspectives for the analysis and design of learning environments. Content analysis of visualizations available online showed that most fail to fully represent the basic principles of natural selection (variation, selection and inheritance). Moreover, the representational potential of visualizations was seldom used to represent threshold concepts such as randomness in origin of variation. Visualizations were also biased to animals as the context of evolution. Similarly, upper-secondary and tertiary students' explanations of natural selection were seldom complete in terms of the basic principles and threshold concepts such as randomness were often lacking. Especially significant was the almost complete lack of randomness in upper-secondary students' explanations. In addition, threshold concepts were context-sensitive across the items used (bacteria, cheetah and salamander), for example spatial scale and randomness was significantly more common in responses to the bacteria item compared to the cheetah and salamander items. Considering the results from these studies, three interactive visualizations were developed (evolution of antibiotic resistance and fur colouration in mice). The visualization design was conducted iteratively

following a Design-Based Research approach and evaluated in classroom settings in secondary and upper-secondary Swedish schools. The results showed that visualizations targeting randomness and genetic level events such as mutations can guide students towards a more scientific conception of natural selection. However, there were differences across the visualizations and student samples. In addition, while students often inferred randomness from the visuals, the results showed that integration of randomness into explanations of natural selection may be challenging. Hence, future research should explore the role of guidance and reflection for students understanding of randomness.

The thesis also discusses the role of students' intuitive conceptions in relation to the use of interactive visualizations and how these preconceptions interact with the presented message. By using the theory of frame semantics, framing effects and conceptual integration, students' issues of achieving an accurate understanding of evolution are discussed in relation to the theory of conceptual change. Implications for teaching and learning natural selection as well as visualization design for learning are also discussed.

Keywords: Evolution, natural selection, visualization, conceptual understanding, conceptual change

Abbreviations

CAQDAS	Computer-Assisted Qualitative Data Analysis Software
COVID-19	Corona Virus Disease 2019
DNA	Deoxyribonucleic acid
EvoVis	Evolution Visualization project
MER	Model of Educational Reconstruction
mRNA	Messenger RNA
NGSS	Next Generation Science Standards
ORI	Open Response Instrument for evolution
PCK	Pedagogical Content Knowledge
RNA	Ribonucleic acid
SARS-CoV-2	Severe Acute Respiratory Syndrome - CoronaVirus – 2
SweFN	Svenskt Frasnät (Swedish FrameNet)
tRNA	Transfer RNA
VR	Vetenskapsrådet (Swedish Research Council)

To my children

Preface

My research interest concerns how we as humans have come to understand nature, both as scientists but also in education. Some phenomena are more puzzling than others but also more revealing. My work on the thesis started as a member of the EvoVis research project, a five-year research grant from The Swedish Research Council. Together with other members of the group, we set out to research the impact of visualizations and abstract concepts (threshold concepts) on evolution understanding. Starting with conceptual understanding and looking at theoretically central content concept of evolution, my interest soon became focused on the concept of randomness in evolution. The reasons for this interest were several, randomness is a central idea of evolution that often seemed overlooked by students but also in teaching material such as textbooks and visualizations. In parallel with my growing interest for students' conceptualization of evolution and randomness, I started to question the focus on described students' explanations of evolution only in terms of the frequencies of isolated concepts. I came into contact with theories of cognitive linguistics and was fortunate to find a graduate course at the Department of Computer Science at the university. Studying some seminal ideas of cognitive linguistics made me realize the parallels between the ideas of conceptual change and conceptual integration or blending. This in turn led me to apply conceptual integration theory and frame semantics as an analytical lens to shed light on how and why students tend to explain evolution in certain recurring ways. This analysis also led to some suggestions on why conceptual change of evolution, especially to conceptions including randomness, might be difficult for learners to achieve. In the concluding section of the thesis, I discuss potential implications of this way of understanding students learning processes of evolution. It is my hope that this thesis might inspire teachers and researchers to further explore the validity of this ideas and also the value of them in teaching and design of learning environments.

Skänninge in February 2021

Acknowledgements

First of all, I would like to express my sincere thanks to my main supervisor *Lena Tibell* for giving me the opportunity to do this thesis. You have tirelessly supported me for many years and always been willing to share your experience, advice and thoughts. I would also like to thank my co-supervisor *Konrad Schönborn* for supporting me during these years. Your sharp thoughts and critical feedback have been very important for my development as a researcher and academic. Next, I would like to express my thanks to *Annika Gullberg* who has read my thesis and papers and given me valuable feedback and discussions during the work. A special thanks also are due to *John Blackwell* for language review of the manuscript,

My thesis would not have been possible without the support from my excellent and helpful colleagues in the Visual Learning and Communication group. You warmly welcomed me as a PhD-student and gave me all imaginable support. It has been a delight and a very learnable journey to work together with all of you. So, thank you Gustav, Gunnar, Daniel, Jörgen, Marta, Henry and Alma!

Some special thanks are due to Gustav Bohlin, Daniel Orraryd and Gunnar Höst. Gustav for including me from the start in the research, for working together on two papers and for sharing your thoughts and supporting me and not the least, being an excellent friend. To Daniel for working together with a paper and for all our interesting discussions and friendship during these years. To Gunnar for being a thoughtful and supporting colleague, working together on two papers, for sharing your knowledge in statistics and for many interesting philosophical discussions.

A special thanks and mention go also to the colleagues at the IPN in Kiel, director Prof. *Ute Harms* and Dr. *Daniela Fiedler*. Daniela for working with me on one of the papers and for our collaborative exchange during the EvoVis project. Ute for welcoming us to the IPN, for constructive and wise feedback and discussions during our meetings and for hosting us on several occasions in Kiel.

I would also like some persons that have been part of our research group on several occasions: *Nalle Jonsson* for supporting me in several ways from valuable and interesting discussions about science but also for together with Lena welcoming me to your home and summer cabin for relief and support during difficult times. *Marie Rådbo* for occasional office company and joyful discussions around science and science education.

Thanks also goes to *Anders Ynnerman* and the MIT-division for hosting me during these years and providing me with opportunities to

develop in several directions professionally. A special mention goes to *Eva Skärblom*, who's administrative support has been the most excellent.

I would also like to thank the colleagues at Norrköping Visualization Center (NVAB): *Sofia Seifarth, Anna Öst, Matilda Stafstedt, Kim Brax* and *Yvonne Olofsson* among others whom I have had the pleasure to work with during these years.

At TEKNAD I would like to thank all of you who have supported me and whom I have worked with in one constellation or other: *Andreas Larsson, Klas Johnsson, Anneli Carlbring, Fredrik Jeppsson, Karin Stolpe, Johanna Frejd, Johanna Andersson, Anna Ericsson* and many others.

A special thanks also goes to biomedical artist and animator *Drew Berry* for collaboration and exchange in the world of biomedical visualization.

Last but certainly not least, the persons who have been my private support and joy in life during these years. My children *Lova, Meja, Elsa, Stina, Signe* and *Hannes* for begin wonderful human beings giving my life meaning and *Josephine*, the mother of my children for supporting both me and the children during the years. My brothers and sisters who have encouraged me and supported me during the ups and downs in life and work. Lastly the person who supported me in every imaginable way during the last and difficult year, *Jenny*.

Sincerely,

Andreas Göransson

Table of Contents

Abstract.....	iii
Abbreviations.....	v
Preface	vii
Acknowledgements.....	ix
Table of Contents.....	xi
Authors contribution to the papers.....	xiv
Sammanfattning på svenska	xvii
Introduction.....	19
Aims and research questions	23
Overall research questions	23
Literature review	25
Evolution – scientific description	25
<i>Origin of variation</i>	26
<i>The relation between genes and traits</i>	28
<i>Changes in allele frequencies</i>	29
<i>Trait loss in natural selection</i>	31
<i>Other evolutionary processes and events</i>	31
<i>Theoretical descriptions of natural selection</i>	31
Evolution – teaching and learning.....	34
<i>Students’ conceptual understanding of evolution</i>	34
<i>Threshold concepts</i>	37
<i>Teaching natural selection</i>	40
Conceptual change and learning evolution.....	41
Visual tools for conceptual understanding	42
<i>Visual representations in Biology – from mesocosm to microcosm</i>	42
<i>Visual representations in biology education</i>	43
Conceptual integration.....	46
<i>Frame semantics</i>	46
<i>Conceptual change, framing and conceptual integration</i>	47
Conceptual framework	49
Overarching framework of the thesis.....	50
The model of educational reconstruction	51
<i>Analysis of science content structure</i>	52
<i>Research on teaching and learning</i>	54
<i>Design and evaluation of learning environments</i>	56
<i>Implementation of MER in relation to the papers</i>	56
Methodology	59
How can learning and understanding be studied?	59
Relation between theoretical perspectives and methods	60
Data collection	61

<i>Samples and contexts</i>	61
<i>Instruments</i>	64
Data analysis	66
<i>Content analysis</i>	66
<i>Statistical analyses</i>	67
<i>Validity</i>	68
<i>Reliability</i>	69
Limitations	70
Design and development of interactive visualizations	71
<i>Paper III Visualizations of antibiotic resistance</i>	71
<i>Paper IV Visualizations of random mutations</i>	75
Ethical considerations	80
Results	81
Research on teaching and learning	81
<i>Paper I</i>	81
<i>Paper II</i>	83
<i>Synthesis of results presented in Papers I and II</i>	87
Implications for design of learning environments	88
Design and evaluation of learning environments	88
<i>Paper III</i>	89
<i>Paper IV</i>	92
<i>Paper V</i>	97
Discussion	99
<i>Representation of natural selection in visual media</i>	99
<i>Threshold concepts in students' explanations</i>	100
<i>The role of interactive visualizations</i>	102
<i>Interaction between preknowledge and taught content</i>	103
Conclusions	105
Implications and outlook	107
Implications for the design of visual media	107
Implications for teaching evolution	107
Future research	109
Presentations	110
References	111
Appendices	125
I. Paper I unpublished data	125
II. Consent form and information to participants	126

Authors contribution to the papers

This thesis is based upon the following papers, which are referred to using the corresponding Latin numerals. For convenience, the studies reported in Papers I-V are sometimes referred to as Studies I-IV.

Paper I

Bohlin, G., Göransson, A., Höst, G. E., & Tibell, L. A. E. (2017). A conceptual characterization of online videos explaining natural selection. Science & Education, 26(7–9), 975–999.

All authors jointly developed the data collection and analysis instrument. I analyzed videos jointly with Gustav Bohlin, participate in data analysis, writing and revising the paper, and produced some of the figures.

Paper II

Göransson, A., Orraryd, D., Fiedler, D., & Tibell, L. A. E. (2020). Conceptual characterization of threshold concepts in student explanations of evolution by natural selection and effects of item context. CBE—Life Sciences Education, 19(1), ar1.

I developed the data collection instrument in cooperation with Daniel Orraryd and Lena Tibell, led the data analysis, performed all statistical analyses and produced all the figures. I analyzed the data in co-operation with Daniela Fiedler, Daniel Orraryd and Lena Tibell. I also wrote most of the paper with suggestions for revisions from the co-authors. The description of the German sample was written by Daniela Fiedler.

Paper III

Bohlin, G., Göransson, A., Höst, G. E., & Tibell, L. A. E. (2017). Insights from introducing natural selection to novices using animations of antibiotic resistance. Journal of Biological Education, 0(0), 1–17.

I researched, designed and developed the interactive animation used in the study, which I designed jointly with the co-authors. I also participated in the data collection and analysis in cooperation with the co-authors, wrote part of the paper and participated in the revisions.

Paper IV

Göransson, A. & Tibell, L.A.E. (2021). Design-based development and evaluation of digital interactive simulations of randomness in natural selection (working title). Submitted.

I conceptualized the study, developed the interactive visualizations, designed the test instruments and methods for data analysis in co-operation with the co-author. I also analyzed the data together with Lena Tibell, and wrote most of the paper with comments and revisions from her.

Paper V

Göransson, A. (2021). Framing evolution – the interplay between students' intuitive conceptions and scientific explanations of evolution. Manuscript in preparation.

Empirically, this paper is based on the same dataset as Paper IV. I developed the test instrument and methods for data analysis. I also analyzed the data and wrote the paper.

Sammanfattning på svenska

Evolutionsteorin förs ofta fram som biologins förenande teori. Vikten av en korrekt och användbar evolutionsförståelse har därför ofta betonats, inte minst för elevers förståelse inom biologins olika delområden men också för att fatta beslut i samhällsfrågor som exempelvis antibiotikaresistens. Många av de centrala delarna av evolutionsteorin är kontraintuitiva eller abstrakta och decennier av forskning har visat att elever har svårigheter att förstå evolutionära processer som mutation och naturligt urval. Representationer såsom visualiseringar, modeller och simuleringar är därför viktiga för att ge elever direkta erfarenheter av evolutionära processer. Även om vikten av visualiseringar är allmänt accepterad inom naturvetenskapsundervisning så är det mindre känt hur visualiseringars utformning specifikt bidrar till att utveckla elevers förståelse av vetenskapliga fenomen såsom evolution. Dessutom har forskningen på elevers evolutionsförståelse till stor del fokuserat på så kallade nyckelbegrepp (variation, selektion och arv) som härletts från vetenskapliga beskrivningar av evolutionsteorin. Dessa begrepp antas vara nödvändiga men också tillräckliga för elevers evolutionsförståelse.

Dock har vikten av icke domänspecifika begrepp kopplade till evolutionsteorin, såsom slump, sannolikhet, spatial och temporala skalor (så kallade tröskelbegrepp), inte undersökts i någon högre grad. Den här avhandlingen använder Model of Educational Reconstruction för att utforska betydelsen av tröskelbegrepp för evolutionsförståelse. Med utgångspunkt i den vetenskapliga beskrivningen och historiken undersöks förekomsten av tröskelbegrepp i befintliga visualiseringar för lärande samt elevers förklaringar för att formulera designprinciper för interaktiva visualiseringar av evolution. Dessutom beskrivs utvecklingen av ett antal interaktiva visualiseringar samt undersökningar av deras potentiella användning i klassrumsmiljöer. Avhandlingen diskuterar även betydelsen av elevers intuitiva föreställningar i relation till användandet av interaktiva visualiseringar och hur dessa föreställningar interagerar med det presenterade budskapet. Genom användning av ramsemantisk teori inklusive "framingeffekter" och "blendteori" diskuteras elevers svårigheter och utveckling av en vetenskaplig evolutionsförståelse i relation till tidigare teorier om begreppsförändring. Konsekvenser av "ramsemantisk teori" och "framingeffekter" i visuella medier diskuteras även i relation till visuell design för lärande.

Introduction

This dissertation addresses secondary and tertiary students' understanding and learning of evolution, seeking to uncover troublesome concepts for students and explore potential strategies for overcoming them. Arguments are advanced that it is important to consider some so-called threshold concepts, especially randomness, in teaching and learning evolution, and that visual representations can be effective means of communication. Empirically, the thesis is largely based on studies (designated Studies I-V, reported in papers designated Papers I-V) in which content analysis of extant visualizations and learners' explanations as well design-based research were used to explore the impact of interactive visualizations on learners' understanding evolution and associated threshold concepts, especially randomness.

During their biology education, students of all age groups encounter visual media of various types or representations with varying interpretational demands. Since many phenomena in biology, perhaps especially evolution, are complex and involve objects and processes that lie outside direct sensory perception, such visuals can play an important role in combination with other modalities such as texts and speech in the classroom. In addition, visual representation is an efficient means of augmenting our cognitive capacities, such as memory or reasoning and hence has substantial potential to enhance learning. However, the factors impacting the effectiveness of different representations of evolution are less well known. Since evolution is known to be a challenging topic for students, it is important to improve our understanding of how visuals can contribute to understanding of evolution.

The research has value for:

- Better understanding of conceptual challenges of evolution for learners
- Better understanding of factors associated with visual media that facilitate or inhibits conceptual change in evolution
- To inform teaching and learning of evolution
- To inform design of interactive visualizations promoting conceptual change and thus scientific understanding

Evolution is often considered the unifying theory of biology (Dobzhansky, 1973; Kalinowski et al., 2010) and evolutionary theory has wide applications in diverse fields, such as medicine, agriculture, conservation biology and environmental science (Hendry et al., 2011). For example, *evolutionary medicine* can help efforts to understand

degenerative diseases and pathogens' evolution, combat pathogens' drug resistance, and improve cancer treatment (Stearns, 2012). *Evolutionary theory* can provide us with understanding not only of the diversity and unity of life on earth but also of our history as the species *Homo sapiens* and our relations with every other living species on the planet.

In addition, public understanding of evolution may be crucial for solving major contemporary social and global ecological problems. Global biodiversity, and consequently the survival of humankind on earth, are threatened by a number of anthropogenic issues such as climate change driven by release of greenhouse gases (Díaz et al., 2019). This threatens the global ecosystem so much that scientists now speak of the sixth mass extinction (Ceballos et al., 2015). While earth has historically experienced climate changes, the rate of the current anthropogenic warming is unprecedented during the last 65 million years (Diffenbaugh & Field, 2013) and exceeds species' ability to adapt due to limited rates of migration and evolution (Meester et al., 2018).

Another environmental issue that requires urgent action is habitat loss and destruction due to anthropogenic activities such as deforestation and agriculture, which is severely straining natural ecosystems (Weiskopf et al., 2020), with potentially catastrophic consequences for the 3.7 billion old life forms on earth and ultimately human society. Understanding *evolution* is crucial to develop better methods of agriculture and forestry and to mitigate the consequences and restore ecosystems (Diffenbaugh & Field, 2013).

At the time of writing of this thesis, the COVID-19 global pandemic, caused by the virus SARS-CoV-2, has severely affected humans and human society. Evolutionary theory is important for understanding the emergence, transmission and changes of the causal agents in such pandemics, and ways to limit their spread and minimize effects of infection (Hendry et al., 2011; Liu et al., 2020).

Understanding *evolution* also provides us with an important set of tools to better understand humans' effects of life on earth and ways to minimize our environmental impact. Thus, it is crucial for protecting and restoring biodiversity, which are key elements of the United Nations' (UN's) global goals for sustainable development (United Nations, 2015). Evolutionary theory has also inspired computer science and engineering, in which so-called evolutionary algorithms are used to solve complex problems (Dasgupta & Michalewicz, 2013), for example design of satellite antennas by NASA.

Thus, an understanding of *evolution* among the general public is desirable to enable citizens to make informed decisions in everyday life. Consequently, evolution is generally regarded as an important topic of biology curricula for various age groups in many countries, and included

for instance in sets of “Big ideas” (Harlen, 2010) and “disciplinary core concepts” (NGSS Lead States, 2013). However, there are also countries where evolution is not part of the official curricula and in some cases it is even explicitly banned as a teaching topic. This problem is not limited to non-western countries, there have been long-standing controversies about the teaching of evolution in countries such as the U.S.A. and U.K. (Harmon, 2011).

Unsurprisingly, given its importance, students’ understanding of evolution has been an active area of science education research for several decades (e.g., Bishop & Anderson, 1990; Brumby, 1979). Moreover, the first article on teaching evolution I found in searches of relevant databases (published 80 years ago) not only deals with the visualization of genes and randomness (Keeler, 1941), but also proposes central concepts of evolution for teaching. The first paper to mention difficulties for students to grasp evolutionary theory was published as early as 1958 (Hunter, 1958).

Nevertheless, contemporary research in science education shows that despite decades of study, evolution still poses significant problems for teachers and students (Smith, 2009a, 2009b). A number of common and recurring misconceptions have been documented (Gregory, 2009), despite advances in assessment of students’ evolutionary knowledge (Anderson et al., 2002; Bishop & Anderson, 1990; Nehm & Schonfeld, 2008). However, most research on students’ conceptual understanding of evolution has focused on their theoretical knowledge of natural selection. Much less attention has been paid to the importance of concepts that are not specifically linked to natural selection (threshold concepts), such as randomness, probability, spatial and temporal scales, all of which are vital for understanding natural selection and evolution (Tibell & Harms, 2017). Such abstract concepts are amenable to visualization for concretization and facilitation of learning and understanding. However, there have been few (if any) studies on the degree that visualizations of evolution effectively exploit the potential to represent such abstract concepts. Neither has research yet provided guidelines for representing such concepts in an intuitive and efficient manner for learners of different age groups.

This thesis addresses some of these issues by presenting examinations of both the content of extant visualizations of evolution and explorative studies of novel visualization designs with particular consideration of these abstract concepts.

Aims and research questions

The overarching aim of the thesis is the question of how we as humans come to understand natural phenomena, more specifically evolution, the central and unifying theory of biology. The purpose is to explore a revised conceptual framework or educational reconstruction of evolution by natural selection with special emphasis on so called threshold concepts. This conceptual framework adds several concepts to the commonly used core or key concepts of natural selection. The exploration focuses on the role of visualization as a teaching medium and students conceptual understanding in relation to visualizations.

Overall research questions

1. How is natural selection represented in terms of conceptual content (key concepts, threshold concepts and misconceptions) in visual media available on the Internet? (Paper I)
2. To what extent do visual media use the potential to represent abstract concepts of evolution such as randomness, probability spatial and temporal scales? (Paper I)
3. How do students use threshold concepts in written explanations of natural selection? (Paper II)
4. How do interactive visualizations focusing on threshold concepts, especially randomness in mutations, affect students understanding of evolution? (Paper III, IV)
5. How are students' explanations of natural selection related to semantic frames and conceptual blends? (Paper V)

Literature review

Evolution – scientific description

“On the origin of species” by Charles Darwin (1859) is regarded today as a science classic of the same magnitude as Isaac Newton’s “Principia” and Carl Linneaus’ “Systema Naturae”. The publication caused a stir in the religious society of Victorian England and although 150 years of biological research has confirmed that Darwin’s theory of natural selection is essentially correct, the ideas in the book are still controversial for some. Evolutionary theory can be regarded as the unifying theory of biology (Dobzhansky, 1973) and provides vital foundations for many sciences aside from biology such as medicine, agriculture, environmental science and biotechnology among others (Hendry et al., 2011). Briefly, evolution can be defined as the process of change in heritable traits over successive generations, and evolutionary theory is a broad term that is essentially encompasses:

- 1) The fundamental tenet that all organisms have a common ancestor, and their subsequent evolution can be verified by clear empirical evidence
- 2) The history of evolution – the dates that lineages split and how each lineage has changed
- 3) The mechanisms involved in evolutionary change (how and why it occurs)

This thesis focuses on the communication and understanding of the history and mechanisms of evolution (especially the latter, component 3 in the list).

A change in heritable traits depends on two major processes (Duret, 2008):

- 1) The *origin of (genetic) variation* by mutations and
- 2) *Changes in allele frequencies* within a population over generations (time).

The second of these processes is sometimes called deviation from Hardy-Weinberg equilibrium, which is based on the general principle that frequencies of alleles and genotypes (concepts described below) in populations remain constant in the absence of other evolutionary influences. A working understanding of the mechanisms of evolution should include at least knowledge of these two major processes (and that they are separate processes), why and how evolutionary change occurs and is

maintained within populations. The following text presents an overview of the essential components of these processes to provide foundations for the educational analysis and reconstruction of evolution.

Natural selection is indisputably an important mechanism of evolution, perhaps the most important. However, other processes, such as mutation, genetic drift, migration and sexual selection (which is regarded as a form of natural selection by some authors and a separate process by others) also make important contributions.

Origin of variation

At the time of writing “On the origin of species” in 1859, variation was central for Darwin’s theory. Without the constant tendency of populations of organisms to produce novel variation, natural selection could not work. While Darwin observed variation in populations and formulated this as one of his central principles, neither science at the time nor Darwin could explain how novel variation occurred. Darwin nurtured the idea that environmental conditions acting on organism’s bodies or reproductive organs generated variation (Winther, 2000).

In contrast, modern evolutionary theory since the Modern Synthesis (reconciliation of Darwinian selection and Mendelian population-oriented genetic theory), attributes (heritable) variation to internal genetic causes rather than external environmental influences (Pigliucci & Kaplan, 2006). It also holds that evolutionary adaptation is caused by external factors of the environment acting upon the existing variation rather than through the environment directly influencing or shaping the traits of a single individual within its lifetime.

However, describing Darwin’s views on the nature of variation as simply caused by environmental changes or conditions, as claimed for example by Winther (2000) is a gross simplification. Importantly, Darwin introduced the idea of chance in biology, especially in origin of variation, as early as 1837 (Hodge, 2016; Johnson, 2015). As Johnson remarks, “*Chance variation may have been an even bigger idea for Darwin than natural selection*” (Johnson, 2015, p. xi). The origin and nature of variation continued to be subject to scientific debate and discussion well into the 20th century, until pioneering experiments such as one by Luria & Delbrück (1943)¹ confirmed that novel variation was indeed random with respect to the environment. That is, adaptive novel mutations do not occur in responsive to selective environmental influences.

¹ Luria and Delbrück’s pioneering experiment showed that bacteria do not evolve resistance to viruses (phages) in response to exposure to the virus, as resistance to the T1 phage (virus) arose randomly in *Escherichia coli* cultures without previous exposure to the virus.

While this pertains to the nature of novel variation or mutations, it was not until the 1950's that science uncovered the nature of the hereditary material and the mechanisms involved in generation of this random variation. The discovery of deoxyribonucleic acid (DNA) by Friedrich Miescher in 1869, with the subsequent confirmation of its role as the hereditary substance by Oswald Avery and finally elucidation of its structure and replication by Watson and Crick in 1953, based on the skilled X-ray analysis of Rosalind Franklin, are key discoveries that finally explained the source of variation and heredity in evolution (Mayr, 1982).

This laid the foundations for molecular biology and paved a way to read the code of evolution directly at the molecular level by DNA sequencing and understand how novel variation enabling natural selection originated through mutations of DNA molecules. Modern evolutionary research is therefore also largely an information science, relying on computer power and algorithms to (for instance) reconstruct evolutionary relationships in the form of phylogenetic trees and the history of living organisms, for example the tree of life. Thus, today's scientific conceptualization of evolution rests firmly upon the science of genetics and the gene is a central concept. A *gene* can be defined as one or more DNA sequences that affect the traits of an organism (Portin & Wilkins, 2017). A gene can exist in one or more versions called *alleles*. In humans who have two copies of each chromosome, an individual can carry one or two versions or alleles of a gene, one on each chromosome in a pair. Organisms that lack sexual reproduction, such as bacteria, often have a single chromosome and hence can only carry one allele or version of a particular gene.

Mechanisms causing genetic variation

Today, we know that all living organisms are composed of one or more cells, and that all contain DNA as the hereditary substance. The DNA contains the genetic makeup of each cell, and is composed of series of four building blocks called nucleotides that form a double helix (spiral). The sequence of the nucleotides determines the so-called genetic code: the set of sequences of three adjacent nucleotides (codons) that specify the amino acids in proteins. These three-letter codes are nearly universal among organisms due to their shared evolutionary origin. During reproduction (for example cell division), DNA is replicated, resulting in two generally identical copies. DNA is replicated by a host of enzymes with different functions such as insertion of nucleotides into a growing strand of DNA and error correction. Despite the amazing fidelity and speed of the replication enzymes, 'copying errors' or mutations are inevitable, so pairs of DNA molecules following replication are not usually perfectly

identical. For example, in the bacterium *Escherichia coli* with a genome size of 5 million base pairs, the general error rate is ca. 10^{-10} mutations per base pair and generation. This might not seem much, but given a typical test tube culture of bacteria with 10^9 cells per ml, an overnight culture in a test tube will contain every possible base-pair mutation across the genome (Milo & Phillips, 2015).

Thus, mutations occur naturally during reproduction and they are the ultimate source of novel genetic variation, but other mechanisms contribute further variation, including (*inter alia*) chromosomal cross-over, random assortment of chromosomes and gene duplication.

It is important to stress that the outcome of these process is random with respect to the selective environment of the organism. Although certain environmental factors may change mutation frequencies, such as ionizing radiation, the mutations are not biased towards adaptive mutations. Thus, variation arises on the genetic or molecular level, and is random with respect to the organism's fitness and future needs.

The relation between genes and traits

Genetic variation can, in turn, cause variation in the traits and behaviors in an organism's physiological anatomical morphological and behavioural traits, which are collectively referred to as its phenotype. Importantly, it is the phenotype that is subject to selection. Since there is a correlation, although not a 1:1 correspondence, between genotype and phenotype, a selection for a phenotype can cause changes in allele frequencies if associated traits have a genotypic basis. Phenotypic variation is not only caused by the genotype but also by environmental factors that influence how genes are expressed. Thus, the same genotype of an organism can give rise to different phenotypes depending on environmental conditions. This is called phenotypic plasticity.

Genes influence traits because, as mentioned, the nucleotide (codon) sequences in DNA encode the amino acid sequence of proteins. During protein synthesis in the cell, the nucleotide sequence of a gene is first 'transcribed' in the form of a messenger RNA (mRNA), molecule, with a nucleotide sequence corresponding to that of the gene. The mRNA sequence is subsequently translated in the form of a protein molecule by a ribosome that 'reads' the genetic code and matches it to transfer-RNA (tRNA) molecules carrying specific amino acids that are linked by the ribosome in the specified sequence. After folding and subsequent modification, the resulting chain of amino acids becomes a functional protein, with properties (governed by its amino acid sequence and folding) that affect cellular processes and ultimately the entire organism. This link between the genetic code and proteins is often called 'The Central Dogma of Biology'. One of the most important consequences of this dogma is

that information flows from genes to proteins (and traits) and not in the opposite direction. Consequently, this refutes historical and intuitive ideas of soft inheritance, i.e., inheritance of acquired traits.

It is worth noting that there is not necessarily a 1:1 to correspondence between genes and traits. Many traits (so called polygenic traits) are under the influence of multiple genes. Also, some genes do not code for proteins but are so-called regulatory sequences, which affect the expression of genes by inhibiting or promoting binding of RNA polymerase (an enzyme that catalyses synthesis of mRNA) and various 'co-factors'.

Changes in allele frequencies

Various mechanisms can cause changes in allele frequencies, which are briefly summarized in the following text:

- **Natural selection:**
 - Positive' or directional selection can increase frequencies of alleles that have positive fitness effects in an organism's environment.
 - Similarly, 'negative' or 'purifying' selection can reduce frequencies of alleles that have negative fitness effects
 - Balancing selection (maintenance of two different versions of an allele) can occur if carriers of both (heterozygotes) have a selective advantage over individuals carrying either single allele (homozygotes). A well-known example of this is sickle-cell anaemia. Homozygotic humans with a mutant haemoglobin gene suffer from this disease, but homozygotes with a 'normal' form are more susceptible to malaria than heterozygotic individuals, who have normal life expectancies and are resistant to the disease malaria. Thus, in environments where malaria is prevalent there is strong selection pressure for heterozygotes.
- **Sexual selection** (sometimes regarded as a form of natural selection) is the evolution of non-adaptive or even negative sexual traits that increase the reproductive success of an individual but would otherwise have neutral or even negative fitness effects. Strong sexual selection may lead to elaborate sexually dimorphic traits. A classic example is the peacock's tail.
- **Gene flow** is a general term for the transfer of alleles between populations, which may occur through migration of individuals from one population to another, and dispersal of seeds or pollen etc. In small populations, gene flow is often an important source

of novel genetic variation. Understanding gene flow is essential for various applications, e.g., conservation of species by establishing habitat corridors between small populations that otherwise would be at risk of extinction due to limited genetic variation.

- **Mutations** that do not affect fitness are not affected by natural selection and can change allele frequencies by chance (see comments below on random fixation of neutral mutations).
- **Genetic drift** refers to changes in allele frequencies due to the random transmission of alleles to the next generation, which can cause loss of some alleles by chance (even if they have positive fitness effects) and fixation of others. Genetic drift is particularly significant in small populations and is also important in applications such as conservation efforts, as it is essential to identify minimum population sizes required to maintain appropriate levels of genetic variation and inbreeding.

In natural selection, allele frequencies tend to increase frequencies of alleles that contribute to greater fitness than other alleles in the population. In contrast, random events alter allele frequencies in genetic drift. Therefore, frequencies of alleles that have positive, negative or no fitness effects can be increased by chance, in stark contrast to the common conceptualization of evolution as a process of progression or improvement. Importantly, the significance of genetic drift is negatively related to populations' size. This has sometimes led to reasoning that genetic drift is a more esoteric and less important phenomenon than natural selection, especially in educational literature and research (Begrow & Nehm, 2012). However, many threatened species typically occur in small, often isolated populations. Moreover, populations of organisms fluctuate, and they may be very abundant sometimes in their evolutionary history and extremely scarce during 'bottlenecks'. Thus, it is important to understand genetic drift in order to reason about issues in conservation biology and protect biological diversity.

In summary, both adaptive (directional) and non-adaptive processes play roles in evolution. The respective importance of these processes has been a source of disagreement and debate among researchers. Until the 1960's, most biologists viewed differences between species as the result of natural selection. Then Motoo Kimura suggested that most evolutionary changes on the molecular level were the result of random fixation of "selectively neutral or very nearly neutral mutations" (Kimura, 1968; Kimura, 1991). While this spurred a debate in the scientific community at the time, more recent data from gene sequencing has largely confirmed Kimura's theory of neutral evolution (Duret, 2008; Kimura, 1991).

Trait loss in natural selection

Often, evolution and natural selection are understood as progressive processes, leading to increasing complexity and more advanced forms of life. However, there are numerous examples of the reduction or loss of traits. A classic example is the loss of eyes and pigment in cave-dwelling organisms such as amphibians and fishes (Protas & Jeffery, 2012). While trait loss superficially may seem different trait gain (evolution of a novel trait), natural selection explains trait loss by the differential survival and reproduction of individuals with different genotypes. In some instances, pleiotropic effects of regulatory genes explain the simultaneous loss and gain in two distinct traits such as improved smell and blindness (Protas & Jeffery, 2012).

Other evolutionary processes and events

Although evolution is often defined as a change in allele frequencies, this definition is too narrow to include other changes that are part of evolution, for example endosymbiosis (e.g., the origin of mitochondria and chloroplasts by uptake of bacterial cells into eukaryotic ancestors) and molecular evolution. In the light of these scientific developments, calls have been raised for an “extended synthesis” of evolutionary theory (Pigliucci et al., 2010), which have yet to reach consensus.

Theoretical descriptions of natural selection

While evolution, as described above, occurs through a range of processes, natural selection has become a central principle in both research and biological education. As already stated, this thesis focuses on the basic understanding of evolution (especially of secondary and upper secondary school students) and the communication of associated concepts. Thus, natural selection is clearly a major process, together with the origin of genetic variation, to address (as further discussed in the chapter on evolution in school curricula). A widely cited description of natural selection is presented in a seminal paper by Lewontin (1970), who recognized three prerequisites:

1. The presence of individual variation
2. Differential fitness (leading to selection)
3. Inheritance

If these prerequisites are present, a change in allele frequency is said to be due to natural selection. However, it is important to stress that natural selection does not explain the presence or origin of novel variation, which

is a requirement for natural selection. Lewontin's formulation, like many others, is intentionally broad and abstract, as a formulation of natural selection should be able to account for as many cases as possible. Thus, the following text briefly discusses these three basic principles.

Individual variation can lead to differential survival in a specific environment (through selection pressure). Variation is present on many levels of organization, e.g., genetic, cell, anatomical or population levels. Variation also exists between populations of a species and within clades, e.g., groups of species within a genus. For natural selection, the variation in a trait within a population is generally regarded as the most relevant.

There can also be differences in reproductive success of individuals that live to reproductive maturity, and hence in the number of viable offspring they contribute to the next generation. As these offspring will carry heritable traits of their parents, this will influence allele frequencies of their population and thus lead to evolution. Therefore, we can extend Lewontin's description of natural selection by including a number of important key concepts of natural selection, and if we include the origin of variation, we end up with the following list of key concepts to explain evolution by natural selection:

- Origin of variation
- Individual variation
- Differential survival
- Reproductive success
- Selection pressure
- Inheritance
- Change in population

Combining the above listed key concepts, we end up with the following description of evolution by natural selection (Figure 1).

In any given reproducing population, variation will arise through mutation of alleles (*origin of variation*). Different allele versions will give rise to *individual variation* within the population. If the individuals vary in a trait affecting survival chances (*differential survival*) and/or their *reproductive success* due to environmental factors (*selection pressure*) and the trait is heritable (*inheritance*), a change in allele frequencies will occur over the generations (*change in population*).

Figure 1. Relations between key concepts of natural selection.

Evolution – teaching and learning

There is substantial consensus among science teachers and scientists about the importance of a basic understanding of evolution. Evolution education has been studied from several perspectives: understanding, assessment, beliefs and religious issues. This thesis focuses on students' conceptual understanding of evolution, more specifically evolution by natural selection, and the role of visual representations. Accordingly, the literature reviewed in the following text focuses on these issues.

Since the 1970's, a growing body of research has attempted to understand the challenges in teaching and learning evolution and several strategies to address the difficulties have been proposed and explored. While there is a rich body of research on evolution education, there have been relatively few reviews and attempts to synthesize the field. Accordingly, calls have been raised more reviews concerning central issues such as student thinking and assessment of evolution knowledge (Ziadie & Andrews, 2018). Notable authors who have reviews parts of this literature include Gregory (2009) and Smith (2009a, 2009b).

Students' conceptual understanding of evolution

Two issues emerge as especially prevalent in the research on students' conceptual understanding of natural selection: the difficulties for students to learn the scientific ideas and the presence of misconceptions. Several causes have been suggested to explain students' difficulties to correctly understand evolution: difficulties in accepting certain facts (such as the historical fact of evolution, the age of the earth, the probabilistic nature of mutations etc.) and exposure to incorrect descriptions of evolution from through various media channels (Aldridge & Dingwall, 2003), popular culture, scientists and instructional material, e.g.;

“Evolutionary adaptation, or simply adaptation, is the adjustment of organisms to their environment in order to improve their chances at survival in that environment.” (Grade 5-8 teaching material, National Geographic²)

While these issues are important, the fact that students still score low on understanding of natural selection and exhibit misconceptions (e.g., Bishop & Anderson, 1990; Nehm & Reilly, 2007), despite weeks or months of instruction, even at the undergraduate level, points to other issues such as cognitive biases (Barnes et al., 2017; Gregory, 2009). In

² <https://www.nationalgeographic.org/encyclopedia/adaptation/>

addition, research has shown that learner's conceptualization of evolution derives from several sources: prior conceptions (scientific and incorrect), religious orientation, biological worldviews and acceptance of evolutionary theory (Demastes, Settlage, et al., 1995; Evans, 2000, 2001).

Teaching and assessment of students' understanding of natural selection has often been based on a content structure derived from theoretical descriptions of natural selection (cf. the previous section), particularly so-called core and key concepts (Anderson et al., 2002; Bishop & Anderson, 1990; Nehm & Reilly, 2007).

Students conceptual understanding according to this content structure has been assessed in several countries and at several education levels (Kuschmierz et al., 2020). The difficulties for students to understand evolution accurately revealed by such studies have often been used in arguments that enrichment learning is insufficient and conceptual change strategies should be used to teach evolution (e.g., Sinatra et al., 2008).

Misconceptions

A prevalent theme in many studies on students' conceptual understanding of natural selection is the presence of misconceptions both before and after instruction. The major documented misconceptions are teleology, anthropomorphism, use and disuse, inheritance of acquired traits, mutations as adaptive response, essentialism and natural selection as an event (Gregory, 2009). Some misconceptions are rooted in epistemological issues such as the citing of religious sources and confusion arising from the misunderstanding of elements of scientific discourse such as law, theory and hypothesis (Smith, 2009a). Accordingly, the importance of understanding the nature of science has been advocated (Smith, 2009a). Another issue pertaining to epistemology is the coherence of students' explanations and the understanding of what counts as a scientific explanation of a phenomenon (Smith, 2009a). Smith therefore raised the importance of studying students' explanations of evolution. In the following text the most common misconceptions are briefly described based on the account given by Gregory (2009):

Teleology. Many studies have found that people of diverse ages and education levels often think that organisms adapt in response to environmental changes or needs (Gregory, 2009; Jakobi, 2010) and this idea often persists during, or even results from, teaching of evolution. The process is often regarded as involving organisms actively influencing or causing their adaptation. Accordingly there is a common belief that it leads to more complex or 'advanced' organisms (Gregory, 2009),

contrary to the numerous examples of evolution leading to the reduction or loss of traits. While the impact of teleology on evolution understanding evolution is debated, it can clearly impair the teaching and learning of evolution.

Anthropomorphism. Anthropomorphic thinking is the application of human-like intentions, desires or feeling to explain actions or processes of non-human organisms and even non-living matter (Kattmann, 2008). Anthropomorphic expressions may be figures of speech, but may also reflect an incorrect understanding, such as evolution being due to human-like planning of a species to survive.

Use and disuse. This thinking pattern presupposes that the use of a feature or trait by an organism somehow leads to it being kept or evolving. A classic example is the webbed foot of a duck, which was historically explained by the duck using its feet to swim leading to development of webbed feet. A variant of this thinking pattern is also found in reasoning about trait loss or reduction, where a trait that is not used will be reduced in the coming generations (Ha & Nehm, 2013).

Inheritance of acquired traits. While this historic idea has been abandoned by science, it continues to surface among students. The idea is that a trait developed through an individual organism's lifestyle or environment will be inherited in the next generation. For example, a frequently running cheetah will develop more muscles and associated features through the activity than cheetahs that run less often, and these traits will be inherited by the offspring. This idea is obviously a very intuitively attractive one since it combines intuitive ideas with inheritance.

Mutations as an adaptive response. This could be thought of as a special form of teleology, where mutations are thought of as the mediating mechanism of change but in a directed or teleological manner. Typically, needs of the organism or the environment is thought to trigger 'needed' genetic changes so that the organisms can adapt (Gregory, 2009; Pope et al., 2017). This clearly conflicts with the scientific, but less intuitive idea that mutations occur at random and hence independently from selection pressure.

Essentialism. Essentialist thinking disregards variation within a population. Rather, a species is conceptualized as have a specific essence that makes it different from other species. Hence, essentialist thinking is at odds with the variation and population-based thinking that is fundamental for the scientific idea of evolution by natural selection. In addition, the conception of a species-specific essence can present an obstacle for accepting and understanding that species change and new species form.

Natural selection as an event. While the scientific description states that natural selection is a gradual process from changes in allele

frequencies in populations, over the course of many generations, students often construe natural selection as an event.

Threshold concepts

The content structure of natural selection in the form of key concepts is based on the theoretical descriptions used in science. However, understanding these concepts and their relations presupposes an understanding of other concepts that are not domain-specific but relate to cognitive factors. Ross et al. (2010) have proposed that randomness, probability, spatial and temporal scales are vital concepts for understanding evolution. Furthermore, they propose that these concepts are so-called threshold concepts. The notion of threshold concepts proposed by Meyer & Land (2003), but obviously with older roots (Tight, 2014), has been used by researchers and educators to identify important conceptual obstacles or thresholds that learners must cross to advance their understanding of a subject. Crossing these thresholds leads to transformed ways of understanding, interpreting or viewing disciplinary phenomena. Thus, threshold concepts represent how domain experts think, and they are characterized by the following features (Cousin, 2006):

Transformative. Grasping a threshold concept entails a conceptual and ontological shift. For example, conceptualizing natural selection as the result of probabilistic selection of random mutations of genes materialized in DNA molecules is both conceptually and ontologically very different from a teleological or need-based view.

Irreversible. A threshold concept once grasped by the learner is unlikely to be forgotten.

Integrative. Threshold concepts are suggested to be integrative in that they reveal previously hidden relations and connections. An evolutionary example is the ability conferred by grasping the idea of spatial scale to realize the commonalities between seemingly very different phenomena such as the evolution of antibiotic resistance in bacteria and cheetahs' running through consideration of the processes from a DNA or gene perspective.

Bounded. Threshold concepts are bounded in a conceptual space, hence they have specialized meaning in discipline discourses, as illustrated by the understanding of adaptation in everyday senses versus and evolutionary biology.

Troublesome. The suggestion is that threshold concepts are linked to troublesome, for example counter-intuitive, knowledge. Thus, learning a threshold concept might be difficult due to this counter-intuitiveness. A prime example in evolutionary biology is the counter-intuitive idea that novel mutations arise randomly rather than through selection pressure or ‘need’.

While the theoretical status of threshold concepts has been debated (Barradell, 2013; Salwén, 2019), they have undeniably served practical and pragmatic functions in research (see Flanagan, 2015 for a comprehensive list of threshold concept studies). The suggestion by Ross et al. (2010) that there are threshold concepts for evolution has an important implication: the content structure and conceptual frameworks used for teaching it cannot solely rely on its theoretical description from the science community. Such content structure often fails to account for students’ issues with the content, hence content structure for instruction is not necessarily identical with the theoretical content of science (Duit et al., 2012). In such cases, identifying the threshold concepts is important as they are major conceptual obstacles for students, and this thesis focuses more on this pragmatic importance than their theoretical status.

Threshold concepts of natural selection

The following section reviews the suggested threshold concepts of natural selection: *randomness*, *probability*, *spatial* and *temporal scales*. These concepts, originating from Ross et al. (2010), have also been treated in a theoretical paper by Tibell & Harms (2017), who suggest that the importance of threshold concepts for understanding natural selection is currently underexplored. In the following text, these concepts are briefly introduced and links to key concepts and misconceptions are indicated. These relations are also clarified in Figure 2.

Randomness is perhaps the most counter-intuitive and transformative idea of evolution by natural selection. Prior to Darwin, probability and chance were not essential components of biological thinking (Mayr, 1991). In this respect, Darwin’s inclusion of chance as a component in a biological theory was ground-breaking, but also provoking for contemporary philosophers and society (Depew, 2016). For many people, randomness implied the ultimate blow to the role of divine influence. Even today, a number of books and papers as well as numerous websites and videos, question the role of randomness in evolution. Intuitively, randomness might be difficult to reconcile with living organisms having structure, function and behavior. Thus, there are both cognitive and

existential obstacles for accepting and understanding that randomness can play an essential role in evolution (Mead & Scott, 2010). In evolution, randomness occurs in *origin of variation*, e.g., by novel mutations, cross-over of chromosomes and chromosomal assortment during meiosis. This implies that the scientific theory of evolution is at odds with teleological thinking. Randomness is also linked to key concepts such as *differential survival* and *reproduction*. For example, survival is partly subject to chance events like meeting a predator or having an accident. Similarly, mating and hence combination of alleles is also influenced by chance. In addition, randomness is an important concept in genetic drift, migration and founder effects among other processes.

Probability. Since several important processes of evolution have chance components, probability is an important concept. Probabilistic thinking is essential to grasp why low probability mutations in fact are very likely, or even inevitable, in a large enough population and number of generations. In addition, both differential survival and reproduction as well as inheritance are probabilistic. This ultimately means that natural selection is not fully predictable and involves diverse probabilities.

Spatial scales. Comprehending evolution and natural selection demands knowledge about processes and objects on different spatial scales, ranging from molecules to entire ecosystems on a global scale (Tibell & Harms, 2017). For example, the genetic composition of a species affects its traits and adaptations to certain geographic areas. For a student this multitude of interlinked processes can be daunting and complex. Understanding genetics is obviously very important in order to reach a working understanding of evolution. Several of the mentioned misconceptions can be linked to spatial scale, perhaps most notably those associated of *origin of variation* and *inheritance* (e.g., *teleology* and *inheritance of acquired traits*).

Temporal scales. Emergence of the geological sciences was a major scientific advance that paved the way for Darwin's theory of natural selection. The notion of deep time, that the age of earth could be measured in millions of years, rather than thousands as adherents of various religions claimed, meant that natural selection was a feasible explanation for the evolution of species. Similarly to spatial scale, understanding evolution and natural selection relies on an understanding of the temporal scales involved. Some evolutionary processes happen in infinitesimally short times, such as mutations in DNA replication, while macroevolutionary changes such as evolution of a new taxon can take millions of years. Understanding the magnitudes of time is also linked to grasping why

seemingly improbable things become very likely, such as a sequence of mutations giving rise to a complex feature such as an eye. Thus, temporal scales are linked to key concepts such as *origin of variation* and *change in population*.

Figure 2. The relation between key concepts, threshold concepts and human perceptual ranges. Figure by the author, from Paper II.

Teaching natural selection

Against the background of the documented difficulties for students to understand evolution correctly and the range of common misconceptions, it is interesting to briefly consider the role of teaching. In the following text, research pertaining mainly to the issues associated with threshold concepts is summarized. Generally, there are few papers explicitly dealing with the teaching of threshold concepts in evolution. For example, only a small number of papers on teaching and instructional strategies concern the origin of variation (Ziadie & Andrews, 2018), while most concern natural selection and human evolution. Similarly, few studies on teaching focus on deep time or molecular evolution (Ziadie & Andrews, 2018). Moreover, most of those few examples concern undergraduate education, and it seems likely that natural selection dominates even more in secondary and upper-secondary education. While many teaching strategies and interventions focus on natural selection, other studies reveal that learning natural selection does not prepare students well for understanding other evolutionary topics (Pugh et al., 2014). The focus on natural selection can also lead to students have difficulties in understanding the difference between evolution and natural selection (Beggrow & Nehm, 2012). In addition, there is a shortage of research on adequate learning goals for evolution education (Ziadie & Andrews, 2018) so the focus on natural selection and key concepts continues to dominate. Hence, the previously mentioned threshold concepts warrant substantial attention in studies of evolution education. Accordingly, there have been calls for research on concepts such as molecular evolution in undergraduate education (Ziadie & Andrews, 2018), in alignment

with the importance of spatial scale as a threshold concept (and associated micro-scale genetic events) noted in the previous section.

Numerous studies on the teaching and assessment of evolution either explicitly or implicitly depart from the theory of conceptual change and constructivist learning theory (Demastes, Good, et al., 1995; E. M. Evans, 2008; Heddy & Sinatra, 2013; Heitz et al., 2010; Jensen & Finley, 1995; Kampourakis & Zogza, 2008; Sinatra et al., 2008). While many studies report successful teaching strategies or tools, misconceptions and incomplete understandings are also often reported (Gregory, 2009; Smith, 2009b). Hence, further research, especially on previously overlooked factors such as threshold concepts, is clearly needed. In the next section, the issue of students' conceptual change in understanding evolution is examined with the aim to identify promising strategies for improvement.

Conceptual change and learning evolution

During the 1970's, science education research began to acknowledge and explore the fact that students often exhibit non-scientific conceptions both before and after instruction (Driver & Easley, 1978). For example, many tend to believe that plants get all or most of their additional mass as they grow from the soil rather from carbon dioxide in the air. Intuitively, this makes perfect sense from our everyday experience of growing plants: they have roots (most plants at least) that penetrate the soil, and if the soil is bad or not replenished with nutrients, they cease to grow or even die. Often, these conceptions are called misconceptions or alternative conceptions in the literature. Misconceptions are conceptions arising from the interaction between everyday conceptions and taught knowledge (Vosniadou, 2013b). In this sense, a misconception is a not entirely successful attempt by a student to make sense of a scientific explanation in terms of her/his own experience (ibid.). Misconceptions are not simple factual errors deriving from wrong information in teaching or teaching materials, but the result of active knowledge construction (Vosniadou, 2013a). The resultant non-scientific conceptions are referred to, sometimes interchangeably, as preconceptions, misconceptions or alternative conceptions (ibid.). Sometimes other terms such as naïve ideas, intuitive ideas or erroneous ideas are also used. Some scholars have tried to differentiate misconceptions, alternative conceptions and preconceptions (Bahar, 2003; Taber, 2014).

There is a general agreement that conceptual change approaches are more effective than traditional approaches to teaching and learning (Duit & Treagust, 2012). Hence, a major theme in science education research has been to elucidate the nature and origin of misconceptions and

also ways to design teaching and learning sequences to facilitate conceptual change. This thesis is mainly concerned with the role of visual representations for facilitating conceptual change, and the following section reviews their importance in biological sciences and potential for facilitating conceptual change.

Visual tools for conceptual understanding

Visualization is a broad term for different external representations. In some instances, the term is also used to refer to internal representations in the mind. In science education, the importance of internal and external representations and their interplay for conceptual understanding and change has been emphasized (Amin & Levirini, 2018, p. 129). External representations or visualizations are often used to collectively refer to pictures, animations, videos and interactive representations such as e.g., simulations (Gilbert, 2008). Simulations can be used to represent a scientific model in a dynamic and accessible manner to facilitate understanding and conceptual change (Smetana & Bell, 2012). In addition, visualizations can provide access to otherwise inaccessible objects and processes (e.g., molecules or cells) which can help students overcome the well-known micro-macro problem (Johnstone, 1991) and help students reason across several levels of organization in biology (Tsui & Treagust, 2013).

Visual representations in Biology – from mesocosm to microcosm

The history of visual media in biology is long and started with drawings and paintings of different plants, animals and their anatomy centuries ago. Still today, illustration of e.g., plants are common in textbooks and field guides, following this tradition. Drawings also quickly became a way to communicate thoughts and findings as well as a means to hypothesize about how ancient life forms that were found in fossil rocks would have looked like.

Many significant scientific advances in biology occurred in parallel with the invention of tools (e.g., the microscope) that allowed researchers to discover objects and processes on small spatial scales. Yet more powerful instruments such as electron microscopes and x-ray crystallography helped reveal life's smallest constituents and processes such as biomolecules, genes and the mechanisms behind heredity and biological variation. Some of the data collected by these novel instruments lie outside the range of visible light and thus demands reconstruction in the form of models, such as a model of DNA or a protein. Visual

representations have also been used as tools to express models by biologists. This way, visuals support both thinking and communication. An iconic example of a visualization of a scientific model is Darwin's notebook sketch of the tree of life (Figure 3). The statistical nature of evolution and lack of sophisticated calculators and computers in the end of the 19th century hampered exploration of mathematical models of natural selection theory. A notable exception is the mechanical model of natural selection developed by Francis Galton (Darwin's cousin) in his attempts to explore and model variation (Stigler, 2010). While Galton was ultimately unsuccessful in modelling natural selection accurately, his contributions in statistics later became useful for evolutionary biology and natural selection, which in the coming decades would be a statistically described (Ariew et al., 2017).

Figure 3. Left: Darwin's tree of life. Notebook of Charles Darwin 1837, Public Domain. Right: The first simulation of natural selection as devised by Galton in 1877, Public Domain. Source: Wellcome Library.

Visual representations in biology education

Visualizations have been traditionally used in education for centuries, with the first example often being Comenius "Didactica Magna, The Great Didactic", published in 1657. Due to the complex nature of life ranging of several magnitudes of scale, biology education often uses visual representations to make things concrete and observable for students. Generally, biology textbooks are rich in visual media and with the ongoing digitalization dynamic visualizations (such as videos or simulations)

are becoming increasingly common. In addition, the increased performance to price ratio of computing devices such as computers, cell phones and tablets has raised an interest in the potential use of these tools in science education.

From a cognitive perspective, we live in the mesocosm (Vollmer, 1984), that is, the part of universe directly available to our senses³. It is claimed that our senses and brains have been adapted over hundreds of millions of years to this mesocosm. However, our brains have an astounding capability of representing and imaging things that we do not have direct access to. Human culture has used external representations for offloading and extending our cognitive apparatus for thousands of years (e.g., cuneiform inscriptions with calculations). This cognitive offloading can free capacity in our working memory which enable us to process more complex information with higher efficiency (Mayer & Moreno, 1998; Paas et al., 2003). Accordingly, visuals can also help organizing complex information and highlight relation among concepts (Tversky, 2011). For example, dynamic visualizations (e.g. animation, simulations and videos) are suggested to help learners understand temporal sequences and causal relations in complex processes (Tversky et al., 2002). In addition, the simultaneous use of images and text can take advantage of dual coding meaning that more information can be processed by using the visual and verbal channel simultaneously (Moreno & Mayer, 1999).

Thus, visualizations offer several important advantages for education. However, there are also challenges. Students must learn the representational culture of the visual discourse in science, e.g., a molecular model representation of DNA or a phylogenetic tree (Anderson et al., 2013). This visual literacy or representational competence is not limited to interpreting visualizations but also encompass the ability to construct and use representations for different purposes. Thus, we should distinguish between the goal of learning to interpret visualizations as a part of scientific culture and methods and using visualizations as a tool for facilitating learning. While visual media can be effective in biology instruction, they are far from unproblematic (Tsui & Treagust, 2013). The scientific correctness of the visual is important to consider but also how learners interpret visualizations and how they affect students' conceptions. Thus, for visualizations to work as knowledge mediators, they must be carefully selected, constructed and used with the learners needs in mind.

³ Not to be confused with the sometimes-used notion of mesoscale in biomedical visualizations, which refers to the scale level between microscopic and atomic levels.

Visual representations of natural selection

While visualizations are used to represent many aspects of evolution and natural selection in teaching, e.g. phylogenetic trees (Schramm et al., 2019), this thesis concerns the representations of threshold concepts and natural selection. Hence, the following section focuses on dynamic visualizations of natural selection and threshold concepts.

According to a paper published in *The American Biology Teacher*, simulations of evolution seems to have existed already in the 1960's (Stebbins & Allen, 1975). However, these were, for natural reasons, not digital but analog, hands-on simulations. Often these simulations were group or class-based activities where students act as organisms, often picking food or resources with tools. The simulations almost exclusively used extant variation within a population and did not simulate the random origin of variation. When digital calculators started to appear in schools, numerical simulations of evolution also began to be used (Forbes, 1978). With the increasing using of computers and subsequently internet in schools during the 1980:s and 1990:s, numerous visualizations and simulations of natural selection appeared with numerous contemporary examples (Appleton & Mackie, 2019).

Despite the plethora of visualizations of natural selection aimed for teaching, relatively few scientific studies on their efficacy have been made. Commonly, simulations are used to represent natural selection, both hands-on and digital. Although there are a range of published papers on the topic, most present the development process (e.g., Hodgson, 2019; Jones & Laughlin, 2010; Paz-y-Miño C. et al., 2011) while fewer investigate students understanding (Brenner et al., 2017; Cardoso et al., 2020; Clarke-Midura et al., 2018a; Dickes & Sengupta, 2013; Herrero et al., 2014; Malone et al., 2018; Pope et al., 2017; Wanford et al., 2018; Xiang, 2011; Yamanoi & Iwasaki, 2015). Notably, many studies were performed with high school level students or above while a few in lower grades (middle school). It should be also be stressed that most studies on dynamic visualizations of evolution and natural selection focus on simulations and not on other formats such as videos or documentaries. While such visualizations often are used in the classrooms by teachers, comprehensive reviews of their conceptual content are currently lacking.

Overall, studies on dynamic visualizations of natural selection targeting secondary school level are scarce. Moreover, only a few of these studies explicitly addressed students understanding of randomness and mutations. Clarke-Midura et al., (2018b) evaluated a redesigned version of their previous natural selection simulation (Pope et al., 2017) attempting to address the role of random mutations for variation. While the

design was successful in reducing the misconception of adaptive mutations, the study did not present any additional evidence of students understanding of randomness.

In summary, studies on the effectiveness of natural selection visualization for instruction are scarce and mostly focus on students above secondary and upper secondary school. In addition, even fewer of the studies have evaluated approaches for communicating randomness, the most troublesome concept for students in connection with natural selection.

Conceptual integration

As mentioned, most assessments of students conceptual understanding of natural selection tend to focus on isolated key concepts or principles, often presented in tandem with misconception frequencies (Nehm & Reilly, 2007; Nehm & Ha, 2011). Fewer studies have focused on the relation between conceptualization of natural selection and language (Nehm et al., 2010) or how key concepts and misconceptions are related (Nehm & Ha, 2011). Cognitive linguistics is a field of research focusing on the relation between language and cognition and emerged as a reaction to shortcomings in formal linguistics. A general assumption of cognitive linguistics is that language provides “*a window into cognitive function*” (Evans, 2019, p. 17), hence the reflection of thought in linguistic expressions is a central tenet. Furthermore, language is seen as a mediator of meaning construction by means of initiating and facilitating a simulation in the listener’s mind. Accordingly, meaning is not transmitted but rather reconstructed, in line with a constructivist perspective on learning. A cognitive linguistics theory purporting to explain how conceptualization occurs is conceptual integration or blending (Turner & Fauconnier, 1995). Conceptual blending builds on frame semantics, a theory of word meaning developed by Charles Fillmore (Fillmore, 1982).

Frame semantics

A semantic frame is derived from a schematized experience that is conceptual and stored in long-term memory (Fillmore, 1982). The frame is represented by different lexical units, i.e., a word. Hence word meaning is related to the structured information of a frame and the meaning of a particular word cannot be understood independently of the related frame. For example, the COMMERCIAL EVENT frame can be represented verbally by word such as *sell* or *buy* (lexical units). Buying or selling, a commercial transaction, typically has a *buyer* and a *seller* exchanging *goods* and *money*. Thus, the word *buy* activates the COMMERCIAL EVENT frame stored in memory and we understand buy in relation to

this background knowledge. Frames also provide a certain perspective and structure on a scene. Hence, a particular scene can be linguistically encoded in several ways which can give rise to different conceptualizations of the same scene. An example is the geographic location where sea and land meet. We can use words such as coast or shore which refer to the same location, coast is typically related to a view from the landside while shore is typically used when we are in a boat. Therefore, the choice of words is related to different conceptualizations of experience, for example how a student understands verbal and/or visual information presented in teaching. A large number of semantic frames have been identified and analysed in the Berkley FrameNet project (Baker et al., 1998) and derivatives in different languages building on FrameNet have also emerged, e.g. the Swedish FrameNet, SweFN (Borin et al., 2010). The catalogued frames and corresponding lexical units are publicly available from databases curated by these projects.

Conceptual change, framing and conceptual integration

Building on the notion of semantic frames, Turner & Fauconnier (1995) developed conceptual integration or blending theory which describes how conceptual elements are combined with frames and also how frames themselves are combined. This theory is often referred to as conceptual blending. A conceptual blend is formed from two different so-called input spaces and results in a blended space which give rise to a novel conceptualization or meaning not present in any of the inputs (ibid.). An classic example of conceptual blending often cited in the literature is ‘This surgeon is a butcher’ (Fauconnier & Turner, 1998). By most of us, this surgeon is understood to be incompetent, although neither a surgeon nor a butcher normally is thought of as incompetent. Hence, the blend gives rise to new meaning. This occurs due to cross-space mappings of counterparts (sharp tools etc.) in the two mental spaces (surgery and butchery) and the fact that the means-ends relation (healing versus butchery) is inherited from only one of the inputs (ibid.)

Conceptual blending underlies several cognitive operations such as conceptual change, metaphors, analogies and counterfactuals (Turner & Fauconnier, 1995). Thus, conceptual blending can be used to describe how and why certain meaning constructions occur when content is taught, for example how scientific ideas taught are framed by preconceptions. Hence, conceptual blends can potentially explain how and why the synthetic conceptions arise when learners try to make sense of taught content. A synthetic conception combines elements of an intuitive idea or preconception with elements from the (taught) scientific idea (Vosniadou, 2019). In terms of conceptual blending theory, such synthetic ideas could be understood as conceptual blends combining

elements and frames from a learner's existing conceptual repertoire and the content encountered in teaching. Thus, frame semantics and conceptual blending offer a cognitive linguistics-based approach to the issue of conceptual change and synthetic conceptions or misconceptions.

Conceptual framework

This chapter describes and discusses the conceptual framework of the thesis. The conceptual framework presents the relation between the different theoretical concepts and constructs of the thesis. According to Maxwell, a conceptual framework is *“the system of concepts, assumptions, expectations, beliefs and theories that supports and informs your research”* (Maxwell, 2012).

The chapter begins with a short overview of science education research in order to situate the thesis in the broader context. Following this, the overarching framework of the thesis is introduced and linked to evolution education research. This framework also serves to structure and relate the different papers of the thesis.

Science education research is a broad and multidisciplinary field drawing inspiration from cognitive science, psychology, anthropology, history and philosophy of science (Treagust et al., 2014). The diversity and intersections between different research paradigms have both been a strength and a weakness of science education research. Inspiration from different fields of inquiry and philosophical stances has clearly enriched science education research but also led to some entrenchment which to some extent has inhibited progress. This has been especially evident within the conceptual change strand of research, where recent attempts have been made to unite the field (Amin & Levrini, 2018).

Students conceptual understanding, which is the focus of the thesis, has mainly been explored within this research tradition called conceptual change (Duit & Treagust, 2012; Lin et al., 2019). Early research in science education revealed that learners often possessed conceptions different from, or incompatible with scientific ideas. This resulted in the idea that science education should aim for a conceptual change within the learner. A seminal paper uniting Kuhn’s idea of paradigm change in science and science education was Strike & Posner (1982). Several criteria for achieving conceptual change were suggested and the change was often understood as the *exchange* of ideas, meaning that students would abandon their old incorrect ideas in the light of the taught content. Despite the research that ensued together with teaching strategies based on the criteria, conceptual change turned out to be a gradual and slow process in many instances. In addition, learners were often found to achieve at best partial conceptual change, often exhibiting hybrid or synthetic conceptions between preconceptions and scientific ideas (Amin et al., 2014).

Conceptual change trace its roots to a constructivist view of learning and cognitive psychology. The constructivist view of learning relies on the assumption that understanding that is (re-)constructed within the

learner as the result of learners' sense-making processes (Amin & Levrini, 2018). This sense-making process occurs because learners are not blank slates but use their current knowledge to understand the new information. As a result of this interaction, students do often end up with a conception that is different from the taught knowledge of information. Accordingly, more recent research following in footsteps of the conceptual change strand has sought to explain how and why students come to an understanding mixing everyday ideas and scientific explanations. Most notable has been three different but somewhat overlapping perspectives on conceptual change: The Knowledge-in-Pieces perspective (diSessa, 1993), Ontological assignments (Chi, 2009) and The Framework theory (Vosniadou & Ioannides, 1998).

This thesis departs from an interest in how learners understand evolution by natural selection. Thus, the science content is a central aspect that has guided the inquiry process. In addition, the constructivist perspective together with conceptual change has been important theoretical lenses for the research design and interpretation of the results. A relatively recent framework for guiding science education research that aligns with a content perspective and a constructivist view of learning is The Model of Educational Reconstruction (MER) described by Duit et al. (2012). This framework also acknowledges the importance for science education research to contribute to the design and evaluation of learning environments, such as e.g., visualizations. In the following, MER is presented and linked to aims of the thesis and the individual papers.

Overarching framework of the thesis

As described above, the overarching framework of the thesis is inspired by a number of extant theoretical frameworks:

- The model of educational reconstruction (MER)
- Conceptual change
- Conceptual integration

The guiding framework for the research design in this thesis has been The *Model of Educational Reconstruction* (Duit et al., 2012; Kattmann et al., 1996). The study of learning and understanding is framed within the conceptual change research tradition with the addition of recent advances in the cognitive sciences, namely cognitive linguistics and conceptual integration in Paper V. MER as an overall framework provides the researcher with a general and flexible framework to conduct science education research and at the same time guides the educational reconstruction of content for design of learning environments, such as

dynamic visualizations. An outcome of MER is also the building of content specific theories of learning and instruction.

This fits well with the overall aim of the thesis; to explore and describe extant visualizations for teaching evolution as well as critically examine these in the light of the theoretical science content and cognitive science in order to develop novel visualizations for learning. In the following, the major components of the framework are briefly summarized and reviewed and their relation to the work of the thesis is presented.

The model of educational reconstruction

The aim of the model is to merge two common strands of research in science education: Focus on science content (*the science-oriented perspective*) and focus on student learning (*the student-oriented perspective*). The following description of MER is based upon the publications by MER aims for a *reconstruction* of science content for teaching with students needs and capabilities in mind. *Reconstruction* should here be understood as the process of rethinking and reflecting upon science content from an educational perspective (Duit et al., 2012). The model has been used in areas of science education such as physics, chemistry, biology, geology and but also in computer science, electrical engineering and science exhibition development (Baalmann et al., 2004; Kersting et al., 2018; Kersting & Steier, 2018; Niebert et al., 2012).

The model has its roots in Klafski's educational analysis (Klafki, 1995) focusing on four key aspects of instructional planning (conventionally referred to as the didactical questions: *Why, what, How* and *By what*. These questions are reflected upon within in the light of students' cognition and attitudes as well as sociocultural preconditions. In MER, this key aspects and questions are structured around three major components (Figure 4):

- 1) *Science content* – key concepts, principles etc.
- 2) *Student and teacher perspectives* on the chosen content.
- 3) *Design and evaluation of learning environments*.

In the following, the three major components of MER are briefly described and related to the previous sections on the scientific content of evolutionary theory, students conceptual understanding and the design of learning environments (in this thesis dynamic visualizations).

Figure 4. The components and their relations in Model of educational reconstruction. (After Duit et. al. 2012).

Analysis of science content structure

The reconstruction of science content is called *elementarization* in MER (Duit et al., 2012). This pertains to transforming scientific knowledge into knowledge for teaching and learning, assuming that the two structures (scientific knowledge and content structure for instruction) serve different needs (science communication versus learning). In MER, special emphasis is put on students' preconceptions for the construction of content for instruction. In short, elementarization considers three different aspects:

- Clarification of subject matter (e.g. the theory of natural selection)
- Educational significance, for example pedagogical goals.
- Misleading terms – e.g. different meaning of words in the science community and everyday life (c.f. the meaning of adaptation in evolutionary theory).

Component (1) is treated by a critical analysis of the evolutionary literature, history of science within evolution and by considering these in the light of current knowledge about student learning. It became clear during the research process that curricula and textbook content (as well as content in visualizations) of evolution to a large extent followed certain traditions, stemming from Darwin's original formulation of evolution and especially natural selection. A critical analysis of the key concepts of evolution as traditionally used for teaching revealed that a number of

essential concepts of evolution were missing. In this thesis, they are denoted *threshold concepts* (see also the section Threshold concepts). Thus, evolution by natural selection can be clarified as subject matter by considering the key concepts and threshold concepts as elementary concepts or ideas that students need to grasp in order to understand natural selection. This thesis mainly concerns secondary and the first year in upper-secondary schools in Sweden. Evolution and natural selection are conventionally taught in grade 8 or 9 in Swedish secondary school. The main reason for focusing on this age group is that Swedish compulsory school ends with grade 9. Hence the natural selection content taught in these grades is often what will be the basis for Swedish citizens knowledge of evolution. In the curriculum for biology in Swedish secondary schools, the following pedagogical goals are mandated (related key concepts in brackets):

Year 7-9

Comparisons from an evolutionary perspective between man and other organisms.

Evolutionary mechanisms and their outcomes, and also heredity (*inheritance*) and the relationship between heredity and the environment.

These quite short and general formulations in the curriculum text is complemented by a commentary material published by the Swedish Agency of Education, where the goals are elaborated upon. The relevant section on evolution states:

“Knowledge development about the origin of genetic variation (*origin of variation*) and how selection occurs by natural selection (*selection pressure, differential survival and reproductive success*). With this knowledge as a foundation, they [students] can explain how changes occur within a population (*change in population*), how speciation occurs and how traits or behaviors evolve.” (Commentary material p 24).⁴

⁴ <https://www.skolverket.se/publikationsserier/kommentarmaterial/2017/kommentarmaterial-till-kursplanen-i-biologi-reviderad-2017>

Research on teaching and learning

The research on teaching and learning reviewed in the previous chapter shows that students in all educational levels face issues with understanding natural selection accurately. Importantly, the dominating content structure for instruction is often based on the basic principles of natural selection (*variation, selection and inheritance*) and elaborated on in the form of key concepts (*origin of variation, individual variation, selection pressure, differential survival, reproductive success, inheritance and change in population*; Tibell & Harms, 2017). The fact that students harbor preconceptions that are at odds with the scientific content and that these preconceptions often persist or develop into misconceptions as the result of evolution instruction presents a challenge for biology education. In the light of these issues, we have reconsidered the content structure for evolution instruction, taking into account common misconceptions and the non-domain specific threshold concepts (*randomness, probability, spatial and temporal scales*). In Table 1, the relations between key concepts, threshold concepts and misconceptions are presented.

The fact that natural selection is a complex process, encompassing several abstract concepts (threshold concepts) makes it amenable to visual representation (see previous chapter, Visual representations in biology education). The literature review of evolution understanding revealed that the reconstruction of evolution for teaching should emphasize that random origin of variation and natural selection are two distinct processes. Studies of historical writings on evolution from Darwin to present day indicates that both Darwin's desire to make natural selection acceptable to contemporary scholars and society as well as the fact that the mechanisms behind inheritance and variation was not known at the time of publication of *On the Origin species* put randomness in the background (Johnson, 2015). Textbooks often devote substantial portion of the evolution chapters to Darwin's life and his discovery of natural selection, while central aspects of evolution such as random origin of variation are less prominent or even omitted. Viewed in terms of MER, the elementarization of evolution has often led to natural selection and its key concepts being in focus but if we take students preknowledge into account, it is revealed that additional scientific concepts should be part of the reconstruction of evolution for teaching. In this thesis these are referred to as threshold concepts.

Table 1 Content structure of natural selection in terms of key and threshold concepts and the relation to common misconceptions.

Key concepts	Relations to threshold concepts	Common misconceptions
<i>Origin of variation</i>	<p><i>Spatial scale:</i> Heritable variation is caused by genetic variation. Linking between different organizational levels.</p> <p><i>Randomness:</i> Novel mutations are random with respect to the selective environment.</p> <p><i>Probability:</i> Probabilities of a certain mutation is low for a single reproductive event but more probable given population size and temporal scales.</p>	Need-based reasoning. Mutations are caused by need or environmental pressure. Mutations are improbable to occur and cannot explain evolution.
<i>Individual variation</i>	<i>Spatial scale:</i> Variation exist on the genetic and individual levels as well as between populations.	Essentialism. All individuals in a population or species are identical.
<i>Differential survival</i>	<i>Randomness and probability</i> are important concepts for a non-deterministic understanding of differential survival.	Only the fittest survive and natural selection as an event.
<i>Reproductive success</i>	<i>Randomness</i> in e.g., mating but also chance and <i>probability</i> in number of offspring.	Reproduction in response to need.
<i>Selection pressure</i>	<i>Spatial and temporal scales:</i> E.g., environmental gradients and changes in environmental factors over time.	Selection pressure causes a need for adaptation or causes novel variation.
<i>Inheritance</i>	<p><i>Spatial scale:</i> Inheritance of observable characters are explained by genes and gene replication.</p> <p><i>Randomness and probability:</i> Chromosomal assortment and cross-over are subject to chance. Inheritance of a trait is related to a certain probability.</p>	Acquired traits are inherited and beneficial traits or genes are always inherited.
<i>Change in population</i>	<p><i>Spatial scale:</i> Change in allele frequencies (genetic level) and in traits (individual and population levels)</p> <p><i>Temporal scale:</i> Changes often occur gradually over the course of many generations. Even small differences in fitness (<i>probability</i> of survival and reproduction) can lead to changes over longer temporal scales.</p>	Natural selection as an event. Populations are always perfectly adapted to their environment and not a continuous process. Only large differences in fitness can cause evolutionary change.

Design and evaluation of learning environments

The third component of MER concerns the design of learning environments (instructional material, learning activities and teaching-learning sequences; Duit et al., 2012). The design should take learners and abilities into account. The main premises underlying the design of learning environments in this thesis is the need to include threshold concepts (the content structure for instruction previously describe) and the need to make these explicit and tangible by means of dynamic and interactive visualizations. In the previous section, Visual representations in biology education, the different pedagogical functions of visualizations are reviewed. It is with these in mind, suggested that visualizations are a suitable and potentially fruitful learning environment for teaching natural selection. The design and evaluation of learning environments can also be used to develop content-oriented theories by inferring students pathways of thinking in connection to the constructed learning environments (Duit et al., 2012).

In paper III and IV, the design and evaluation of three different learning environments for teaching natural selection with an emphasis on threshold concepts is described. In Paper V, students' conceptual integration of key and threshold concepts with preconceptions of evolution is in focus. The analytical lens used in Paper V is conceptual integration or blending (Fauconnier & Turner, 1998). More specifically, how students understanding of the causal structure of natural selection (random variation and probabilistic selection) can be described in terms of conceptual integration and frame semantics theory.

Science education research can be described according to a continuum between basic and applied research. (Duit et al., 2012) argue that science content must be reconstructed according to an educational perspective with the needs of education practice in mind. As a means to bridge the gap between basic research and applied research, Design Based Research (DBR) (Cobb et al., 2003) has been developed. Design based research provides an overall framework for iterative design of learning environments. In this thesis, students' conceptual understanding of the designed learning environments (i.e., interactive visualizations) is assessed by students' written explanations, reasoning and discussion during interaction with the visualizations.

Implementation of MER in relation to the papers

In Figure 5, the relation between the three components of MER and the papers and sections of the thesis is presented. The clarification and analysis of the scientific content is found in the Literature review section of the thesis. The proposed content structure for instruction (key and

threshold concepts) is used as an analytical framework and basis for the design of learning environments described in Paper III and IV. In Paper II, the extant literature on students' conceptions of natural selection is complemented with an empirical investigation of students use of threshold concepts in explanations of natural selection. The results from Paper II were used to complement the content structure for the design and evaluation of novel learning environments (interactive visualizations) in Paper III and IV. In paper V, the analysis of students' perspective (conceptions) is iterated based on the empirical results from Paper IV. Thus, Paper V is mainly related to MER Component 2.

Figure 5. The relation between components of MER and the papers of the thesis.

Methodology

This chapter provides a description of the different methods for data collection and data analysis that have been used in the thesis. It also discusses potential limitations of the chosen methods.

How can learning and understanding be studied?

Many different perspectives and theories can be applied to study the learning, such as cognitive, sociocultural or psychological. The plethora of different ways to define and study learning has turned science education into a multi-disciplinary and constantly evolving field of research encompassing different philosophical and theoretical approaches or traditions, each with its own merits and drawbacks. Commonly, the science education research stems from an interest from domain experts and teachers fostered within the respective scientific field such as biology, chemistry or physics. Over time, science education research has attracted other scientific fields such as psychology, cognitive science and social sciences. The consequences are that science education research of today is a multi-faceted and complex field employing sometimes incompatible but often complementary theories and perspectives (Treagust et al., 2014). The choice of a certain theoretical approach is necessary for the rigor of the scientific work, but this does not necessary imply that this theoretical perspective claims to encompass all important aspects or deny other approaches as less interesting or suitable. This thesis builds primarily on the conceptual change tradition of science education research with a cognitive perspective as the guiding lens of inquiry.

Learning can be studied by quantitative or qualitative methods, or a combination (mixed methods; Treagust et al., 2014). Typically, qualitative methods are used to understand the cognitive processes when students engage with science content (Erickson, 2012) while quantitative methods can inform about the prevalence and significance of a measure in a sample. Each approach has its strengths and weakness and hence the call has been raised for so called mixed-methods research in social sciences (Robson, 2011).

As conceptual understanding is constrained to inner mental processes, the researcher must rely on measures derived from different instruments and the behaviour these elicits (Reeves & Marbach-Ad, 2016). Using students written explanations to test items is a common method to gather information about their conceptual understanding of a domain. While test instruments often contain open ended items, the answers to such items are often scored according to the presence of concepts in isolation (e.g., Nehm & Reilly, 2007). Several papers have previously

studied students use of concepts in written responses to open-ended items (Bishop & Anderson, 1990; Nehm et al., 2010; Nehm & Reilly, 2007; Nehm & Ha, 2011; Nehm & Schonfeld, 2008) and oral explanations of evolution (Nehm & Schonfeld, 2008).

Relation between theoretical perspectives and methods

This thesis departs from the theory of conceptual change. Research on conceptual change usually apply a constructivist epistemology (Amin et al., 2014; Duit & Treagust, 2012). The assumption is that conceptual understanding is the result of personal (and social) knowledge construction occurring when learners face a novel situation (Duit & Treagust, 2012; Vosniadou, 2019). The conceptual change perspective, not exchange, stresses knowledge construction as an interaction between preconceptions and taught content. Hence, students are suggested to attempt to make meaning from novel content with the repertoire of preconceptions. Accordingly, partial conceptual change is to be expected and seen as normal, typically exhibited as mixed or synthetic conceptions between intuitive ideas and taught scientific ideas.

For assessing students content knowledge in relation to a framework of scientific ideas, deductive content analysis is suitable (Hsieh & Shannon, 2005). This was used in Paper II to describe students use of threshold and key concepts of natural selection in relation to commonly used test items.

However, capturing students learning processes and attempts to construct meaningful explanations with only the presence or absence of scientific concepts is less informative if we assume that pre-knowledge or preconceptions are important for how students construe taught content. Accordingly, the analysis was widened in Paper III, IV and V to include more inductive content analysis in order to capture the different ideas students expressed when interacting with visualizations of natural selection.

Describing concept use in isolation insufficient to capture students learning process (Catley et al., 2005). The data collection and analyses hence should capture not only correct or incorrect ideas but also inform about students attempts to construct an understanding. Multiple-choice items are not suitable for uncovering students' individual ideas since these might not fall into the predetermined answer alternatives (Nehm & Schonfeld, 2008). Synthetic and mixed conceptions are common as a result of science teaching (Demastes, Settlage, et al., 1995; Nehm & Ha, 2011). To capture these, open-ended items is a feasible instrument since it gives students an opportunity to freely express their thinking.

Similarly, group discussions and tasks can inform about student thinking and learning when working with external representations such as visualizations.

Data collection

The thesis builds on five datasets. Dataset 1-3 is used for paper 1-3 and dataset 4 and 5 are used in paper 4.

- Dataset 1 consists of students’ explanations of natural selection collected from upper secondary and university students in Sweden and Germany.
- Dataset 2 consists of videos sampled from the internet.
- Dataset 3 consists of pre and post-tests and written answers and video recordings from group discussions.
- Dataset 4 consists of pre and post-tests and written answers from group discussions.

Table 1. Datasets, their source and studies in thesis.

Dataset	Source	Study
Dataset 1	Written answers	I
Dataset 2	Online videos	II
Dataset 3	Written answers and video recordings	III
Dataset 4	Written answers	IV,V

In the following sections, sampling, contexts, data collection instruments used, and data handling is described.

Samples and contexts

Paper I. Publicly available visualizations were collected from the Internet. The sampling was limited to videos for practical reasons and comparability issues. The sampling approach chosen was intended to mimic frequently used search strategies on the Internet. Searches were performed via Google Video and YouTube and combined with searches on the result pages (Ortiz-Cordova et al., 2015). When a resource of fitting the criteria for inclusion (see below) was found on a web site, further searches were performed on this web site. Videos found via this approach were also considered for inclusion in the sample. This approach was completed with targeted searches on specific platforms (e.g., HHMI

Biointeractive, evolutionfilmfestival.org – see Paper II Methods for a comprehensive list). Searches were performed by using keywords such as “*evolution*”, “*natural selection*” combined with “*video*”, “*animation*”, “*visualization*” and “*explanation*”. Search results were collected in a database according to certain inclusion criteria to ensure that videos were depicted biological evolution, non-creationist and had an explanatory approach towards evolution. The final sample consisted of 60 videos.

Paper II. Participants were sampled from Sweden and Germany universities. In Sweden, data was also collected from upper secondary school classes. The latter subsample was not used in the published paper but is presented in the results section of the thesis. Participation was voluntary and anonymous. For the Swedish subsamples, no incentives were given. In Germany, participants were recruited to the online survey via student councils web sites. Here, all volunteers taking the survey were offered participation in a lottery for 10 vouchers (each worth approx. US\$54). In addition, 32 biology students were recruited from Kiel University, each receiving US\$32 for participation.

The Swedish sample consisted of 38 undergraduate students (average age 23,7 years) from a south Sweden university. The students attended primary teacher education (n=4) and chemistry- or biology-oriented programs (n=34). All students at this level had previously received teaching about evolution in their preceding upper secondary education and the students from the biology-oriented education program had received an additional introductory biology course covering the basics of evolutionary theory.

The German sample consisted of 209 biology students from 21 German universities (average age 23 years) of whom 97 where biology majors, 71 undergraduates (bachelors’ degree) and 26 graduate students (masters’ degree). The remaining 112 students were preservice biology teacher students of whom 52 at undergraduate level, 42 at graduate level and 18 at the basic level courses. All students at this level in Germany are normally exposed to teaching of the basic evolutionary theory and mechanisms (see Paper III, Appendix A for further details).

In *Paper III*, 32 Swedish 8th grade students (16 females) aged 13-14 years were recruited on a voluntarily basis. No incentives were given for participation and informed consent for participation was given by the students and their parents prior to the beginning of the data collection. All the students belonged to the same class and the data collection was done in their regular classrooms. The data collection was performed as a part of their regular biology course, before the course unit on evolution (which is usually given in grade 9 in Swedish secondary school). Striving for a natural classroom context, students were assigned by the teacher

into groups of 3-5 students forming a total of eight groups. Five students not consenting to be video-taped were assigned to the same group. From this group, only written answers were collected. Data was collected anonymously and in order to track students and groups and matching them with their pre- and post-test answers, a tracking number was used for each individual. Before the data collection begun, the researchers and teacher gave a short oral introduction to the study and natural selection including an explanation of the relation between DNA, genes and inheritance.

The dataset for *Paper IV* and *V* was collected on three different occasions. All recruitment was on a voluntarily and non-incentive basis, using informed consent from the participants. All data was collected anonymously and tracking numbers were used for each student in order to relate different answer sheets to the correct student.

For Visualization 1 (antibiotic resistance), data was collected from two different upper secondary school classes. The first class (11 students, 5 female, average 16,0 years, $SD=0$) attended a small independent school located in a larger town in south east Sweden. The second class (19 students, 11 female, average 16,6 years, $SD=0,87$) attended a public school in a south Sweden mid-sized town. Both classes attended their first year in the national Natural Sciences program and the data collection was performed in a classroom setting during the Biology 1 course (Skolverket, 2013). In both instances, data collection was performed before the start of their course unit on evolution but after the genetics unit. During the data collection, the regular class teachers in biology participated. A general introduction was given to the topic of antibiotic resistance by the researchers and a short movie introduction produced by the first author was shown. The movie lasted approximately 1 min and gave a brief historical review of antibiotics, microbial infections, antibiotic resistance and its consequences for public health. For the group phase, students were assigned to groups of 3-4 students by the teacher. In total 8 groups were formed.

For Visualization 2, data was collected in a secondary public and compulsory school class in grade 9 (23 students, 11 female, average age=15.1 years, $SD=0,34$ years). The school was located in a south east Sweden mid-sized town. During data collection, two regular class teachers and two researchers participated. Before the data collection, the first author gave a short oral introduction complemented with images to the context of the visualization. Basic relations between DNA, mutations, genes and phenotype were reviewed as well as basic facts about the deer mice and their ecology. The slide show was also available with a summary transcript in the interactive visualization for reference. For the group

phase, students were assigned into groups with the assistance of the teachers (4 students per group).

Instruments

Survey

In Paper II a survey consisting of a subsection of commonly used open-ended test items was used. The items were sourced from the Open Response Instrument (ORI) for evolution, originally used in Bishop & Anderson (1990) and further by Nehm & Reilly (2007). Items referring to concrete evolutionary examples and contexts were used and items asking about general problems such as the definition of natural selection were excluded, since we were interested in also comparing item contexts. The resulting items matching the inclusion criteria were three:

Table 2. Items used in the data collection for Paper II.

Item	Question
1 Bacteria	Explain why some bacteria have evolved a resistance to anti- biotics (that is, the antibiotics no longer kill the bacteria).
2 Cheetah	Cheetahs (large African cats) are able to run faster than 60 miles per hour when chasing prey. How would a biologist explain how the ability to run fast evolved in cheetahs, assuming their ancestors could run only 20 miles per hour?
3 Salamander	Cave salamanders (amphibian animals) are blind (they have eyes that are not functional). How would a biologist explain how blind cave salamanders evolved from ancestors that could see?

The survey was comprised of the above-mentioned items together with demographic variables (age, education level and program). The survey was distributed online to students volunteering to participating (please refer to “Sampling and contexts” below for further details).

Pre and post-tests

Pre and post-tests were used in the data collection in Paper III, IV and V. In Paper III, the pre-test was comprised of one multiple-choice item which also occurred on the post-test. On the post-tests, students were also asked to motivate their choice of answer. One open-ended item

intended to measure transfer to a mammal context (giraffe) was also added to the post-test.

Table 3. The items used in the pre- and post-tests in Paper III.

Pre and post item	<p>Which of the following best describes how bacteria develop resistance to antibiotics?</p> <ul style="list-style-type: none"> A. Bacteria always try to develop resistance when they are exposed to antibiotics. So, those that succeed will be protected the next time. B. Bacteria can become resistant if they infect a person who is already resistant because he or she has used too much antibiotics previously. C. Bacteria that have become resistant through <i>random</i> mutations can survive and spread when antibiotics kill non-resistant bacteria. D. Bacteria that have made a person ill will develop resistance if the person does not finish his or her course of treatment.
Transfer item (Post-test)	<p>Earlier generations of giraffes did not have as long necks as those found today. Try to describe similarities and differences between the neck-development of giraffes and bacterial development of antibiotic resistance.</p>

For Paper IV and V, dataset IV was used. This data set included pre and post-test items mainly sourced from the ORI (Nehm & Reilly, 2007) and that had been used in Paper II: The Bacteria item and the Cheetah item. While these items offer an opportunity for comparison ours and others publication, the contexts are different from Visualization 2 (evolution of fur colour). In order to compensate for this, a new item was constructed for the post-test in Visualization 2, This item asked: “*Polar bears and brown bears have a common origin from a species of brown bears. Explain how the light-coloured fur of today’s polar bears has evolved. Please be as complete as possible.*”

The responses to the pre- and post-test items in Paper IV were analysed with content analysis (see further below).

Group discussions

In the collection of Datasets 3 and 4 (Papers III, IV), group tasks and questions were used to elicit students’ thinking and reasoning when working with the visualizations. The tasks were intended to guide the students in how to use the visualizations and the questions were

intended to support students' reflection and also elicit their thinking about the visualization content.

In Paper III, the questions aimed to probe students' ability to make predictions about evolution of antibiotic resistance, explain the origin of resistance and one item evaluating the visualization. Each question was supplied on a separate sheet of paper following the sequence of the visualization. The group discussions were video recorded and subsequently transcribed verbatim.

In Paper IV, student groups discussed and answered questions posed on the worksheets supplied with the visualizations. The instructions and questions are found in the section Design and development of interactive visualizations.

Data analysis

Content analysis

To analyze and characterize the conceptual content of written answers (Paper II, III, IV), online videos (Paper I) and video recordings (Paper III), content analysis was used. Content analysis is a systematic method to study documents and other communication artifacts such as pictures and videos (Krippendorff, 2013). It is useful for identifying and quantifying the presence of concepts in communication artifacts such as written text, images and verbal data. In addition, it is used to find patterns and correlations in concept use and how concepts are communicated as well as revealing communication differences (e.g. concept use) in different contexts (c.f. Paper II) (Gale et al., 2013; Krippendorff, 2013). The content analysis used in this thesis mainly follows the so-called framework method which means that an analytical framework is used to code and categorize the data. The analytical framework can be developed deductively, inductively or both, thus making it suitable for both capturing data to answer the research questions but at the same time flexible enough to capture unexpected phenomena (Gale et al., 2013).

The analytical framework developed was used to develop a code book or criteria catalogue, initially for content analysis of online videos. The catalogue was inspired by previous work by (Bishop & Anderson, 1990; Nehm et al., 2010; Nehm & Reilly, 2007) with additions of threshold concepts and misconceptions. This criteria catalogue was later modified and used for analysis of students written explanations of natural selection in paper 2.

In paper 3 which was largely an explorative study, inductive and thematic coding was used in order to capture students' discussions about the visualization as broadly as possible.

In paper 4, the criteria catalogue from paper 2 was reused with the addition of model type of the explanations. The different model types was coded deductively, inspired by Vosniadou (2019), in order to capture the degree of conceptual change after using the visualizations. In addition, students' causal explanations of natural selection were coded by an inductive approach to capture the range of different conceptions as holistically as possible.

In Paper 5, deductive content analysis was used to identify semantic frames in the student explanations in a subset of Dataset 4. The subset used were explanations from the pre and post-tests about antibiotic resistance collected during students use of Visualization 1 in Paper 4. Lexical units (words) were used to identify the organizing semantic frame of the explanations. By using the Swedish FrameNet (Borin et al., 2010) and Berkley FrameNet (Baker et al., 1998) databases, frames identified by searching for lexical units and referencing the corresponding frame structure. Conceptual blends were identified as synthetic models combining elements of students everyday experiences and the taught scientific model (Hrepic et al., 2010).

Statistical analyses

In Paper I, the code frequencies for each variable were calculated. Cluster analysis was used to explore patterns and relations within the data (Everitt et al., 2011). Two cluster analyses were performed, one which analyzed relations among the variables and one which analyzed groupings of videos. Variables with high or low occurrences were excluded since they are less informative about differences within the data (Hamid et al., 2010; Islam et al., 2014). Hierarchical clustering based on group average linkage with Yule's Q as similarity measure was performed. The algorithm starting with initial variables as clusters and yields successively larger clusters until a single cluster remains. The number of final clusters were determined by the agglomerative coefficient.

In Paper II, the resulting binary variables were analysed with descriptive and inferential statistics in IBM SPSS 24. The frequencies of concepts were summarized for each item. To identify statistically significant differences between items, Cochran's Q was used. This test is suitable for comparison of binary variables in related samples, such as pass/fail frequencies on test items (Siegel & Castellan, 1988). Where Cochran's Q identified a significant difference due to item context, Dunn's post-hoc test (non-parametric) with Bonferroni-correction for

multiple comparisons was used to identify which items were significantly different.

In Paper III, the frequencies of responses for the pre- and post-test item were compared using McNemar test to find statistically significant changes ($p < 0.05$). McNemar tests are appropriate for paired nominal data.

Paper IV only used descriptive statistics of code frequencies and no inferential statistics.

For the statistical methods used for inter-rater reliability calculations, please see Reliability below.

Validity

In general, validity can be described as a construct that describes to what degree selected research methods capture the intended phenomena. In content analysis, validity can be thought of the degree that the code system represents the phenomenon (Gale et al., 2013; Krippendorff, 2013).

Content validity of the test items used and the criteria for coding of variables (key and threshold concepts) was established by reviewing of an expert panel consisting of subject domain experts (Reeves & Marbach-Ad, 2016) and science education researchers. In addition, the validity of the test items reused from previous publications had been established in prior research by using similar methods, see (Bishop & Anderson, 1990; Nehm et al., 2010; Nehm & Reilly, 2007). The validity of the method to code student answers of natural selection has been established by prior studies comparing the coding of open items with interviews and established measures of natural selection knowledge (Nehm & Schonfeld, 2008). Adding to the validity of the items used, the items were found to elicit all concepts from the analytical framework, an important measure of test-item validity (Reeves & Marbach-Ad, 2016). The fact that key concepts of natural selection were detected for a majority of the responses is also an indicator that although threshold concepts were usually present to lower degrees, the items in fact mirrors students' explanations of natural selection.

Content analysis has some major advantages in being unobtrusive, transparent and replicable. However, a limitation is the reduction of data to codes that are decontextualized. In order to minimize the effects of reduction, codes were always traceable to original data by means of CAQDAS software (MaxQDA) and Microsoft Excel to keep context in analysis. In addition to quantify the presence of concepts using the analytical framework, concept context and expression was coded inductively in Paper II. For Paper III, we employed two additional analyses to capture the relation between concepts: analysis of models of conceptual

change (deductive) and holistic description of the causal models used by students (inductively). Comparisons of these different coding approaches adds to the validity of the conclusions (see Paper III).

Reliability

According to (Robson, 2011), reliability is the “*consistency over time and with different observers*”. In content analysis this translates to the consistency of coding as well as the agreement of coding between the analysts or raters (Krippendorff, 2013). To ensure reliability in our analyses, two major measures were taken: The use of a code book or code memo and the use of two or more raters with the subsequent analysis of inter-rater reliability (Paper I and II). In Paper IV a consensus method was used by the researchers, where successive coding rounds were performed, and code definitions were discussed until the analysts agreed.

For Paper I, reliability of the coding was established using inter-rater reliability calculations. The 60 videos in total were analysed by two independent raters with an overlap of 33 videos. The inter-rater reliability was calculated as Krippendorff's alpha (Hayes & Krippendorff, 2007) where an alpha-value of 0.70 or higher was used as an acceptable level of agreement. Most variables reached above 0.70 (see Paper I appendix for full list). However, Krippendorff's alpha require higher agreement for variables that are present or absent to a large extent in the data. For example, a variable reaching 97% agreement (identical coding by the raters) might still have an alpha close to zero. Therefore, to meaningfully assess reliability, the presence or absence of a variable together with the amount of agreement should also be regarded.

In Study II, representing a similar type of data material with several binary variables coded by independent raters, reliability was calculated by using Guilford's G (Holley & Guilford, 1964). Guilford's G performs more consistently for variables with low occurrences (Xu & Lorber, 2014). If a variable did not reach the desired threshold ($G > 0.7$), definitions in the code book were adjusted and recoding were performed of these variables across the entire sample. Final reliabilities are presented in Paper II, Appendix C.

For Paper III and IV, being mainly a qualitative or mixed methods studies, a somewhat different approach to reliability was used (Robson, 2011). In qualitative research reliability and validity is often replaced with the concepts of trustworthiness and generalizability (Robson, 2011). Trustworthiness can be operationalized as transparency and consistency (Gale et al., 2013). To ensure transparency, rich coding examples are provided in Paper III together with descriptions of the data collection process, sample context and participants as well as analyses. A threat to trustworthiness in qualitative research is observer bias, especially if a

single researcher performs the coding. To minimize this, the coding was performed by two or more independent raters using an iterative approach to coding. A code memo was constructed and used for constant comparison to avoid code drifting, thus ensuring consistency over time. After each successive round of coding, categories were discussed and systematically checked against the data and the research questions.

Limitations

A number of limitations should be taken into account when interpreting the conclusions of the thesis. The first limitation is sample size and context. For Paper I, an obvious limitation is the sample size of 60 videos. Although we employed a sampling strategy intended to mimic naturalistic search behavior, the vast amount of available material on the Internet means that there is likely some videos that are more complete in conceptual content, as well as less, than the analyzed sample. However, we believe that the sample size is large enough to draw the conclusion that many videos do not represent abstract content or concepts of natural selection (e.g., the threshold concepts). This should be enough to warrant the caution of teachers (and students) when searching for videos purporting to explain natural selection.

For Paper II-V, sample size is a limitation. Neither of the samples were drawn with the intention to statistically represent the whole population of students in the respective grades. Therefore, generalization to populations outside the samples used should be made with some caution. There might exist populations performing better or worse on natural selection knowledge. In addition, the visualizations used in Paper III, IV and V might have a different impact in populations that are substantially different than our samples, for example in terms of pre-knowledge of genetics and evolution.

Another limitation for Paper III-V is the lack of control groups that did not use the visualizations. Typically, control groups are not used in early explorative studies of learning designs but are performed at later stages in the evaluation, after refinements have been done to the design (Cobb et al., 2003). This limitation means that the conclusion cannot be drawn that visualizations are superior to other means of instruction.

Design and development of interactive visualizations

This section provides an overview of the design of the interactive visualizations used in Paper III, IV and V.

Paper III Visualizations of antibiotic resistance

During the design of the visualization, we focused on the threshold concepts of randomness, probability, spatial scale and temporal scales. The context chosen was evolution of antibiotic resistance in bacteria. A storyboard was developed by the authors with a linear structure intended to guide students step by step through the process when antibiotic resistance evolves. A key feature of the design is the linking between the molecular and genetic scale events and processes and population scale.

The visualization was developed by using a set of multimedia tools to produce and combine the different sections. Some of the animations were developed using the open-source software Blender3D which enables animation and scale transitions seamlessly. Other content was illustrated with Adobe Illustrator and Photoshop. The protein and DNA structures depicted were imported from the Protein Data Bank. After visualization of the structures in UCSF Chimera, Python scripts were developed to generate animations of DNA-replication. Audio narration was recorded and pre-processed. Finally, the content was brought together in the multimedia authoring tool Adobe Flash, interactive assets were added for navigation and the complete module was exported for use in web browsers supporting Flash.

Brief description of the visualization

The final interactive visualization was comprised of seven linked animations (See A-F in Figure 6 and Figure 7 to Figure 11 below) which can be navigated by the students. The introduction aims to set the context for the student and after this, evolutionary concepts are introduced in a sequential order. The introductory screen briefly explains antibiotics and instructions for navigating the visualization (A). In scene B, the size of bacteria and their inner components such as DNA is introduced by the means of a zoom-in animation and scale bars in order to help students appreciate the microscopic scale of bacteria and that invisible components such as DNA are important. In (C) a schematic picture of a bacterium is presented together with links to animations of DNA-replication (D) and mutation (E). Following this, cultivation of bacteria samples in a laboratory is introduced (F) together with an animation depicting the growth rate and mutation rate in the bacteria culture. Finally, three agar plates are depicted in (G) of whom two contain antibiotics. The students

are asked to predict the appearance of the colonies before the plates are incubated and the final result is presented on the screen.

Figure 6. The overall structure of the sections of the interactive visualization in Paper III. Letters A-F are referenced in the text description of the visualization above.

The sections of the interactive visualization are shown in screenshots below.

Figure 7. Screenshot showing the laboratory environment before zoom-in.

Figure 8. Section (B), inside the test tube during zoom-in with scale bars indicating the size of bacteria.

Figure 9. Section (C) where students can start animations of DNA-replication and mutation.

Figure 10. Section (D), animation of DNA replication

Figure 11. Section (G) showing agar plates with bacteria after incubation. The two plates to the right contain antibiotics and a few resistant colonies are growing on them.

Paper IV Visualizations of random mutations

In Paper 4, two different interactive visualizations were developed. Visualization 1 was a simulation of how random mutations can give rise to antibiotic resistance. Visualization 2 focused on the role of random mutations in evolution of fur color in deer mice. The design of each visualization is briefly presented below.

Visualization 1

In order to focus more explicitly on the randomness of mutations compared to the visualization in Paper III, a new design was developed. The aim was to explore a comprehensive manner to show randomness. In doing this, the first part of the visualization contains an interactive simulation of test tubes with overnight cultures of bacteria. Test tubes containing a mutation for antibiotic resistance is denoted with a red circle (Figure 12). Students can generate new batches of overnight cultures by the click of a button. Each time the button is clicked, a new row of test tubes with overnight cultures appears on the screen, below the previous. A random function in the simulation decides which test tube(s) contain a mutation. The idea was that the repeated generation of bacteria cultures would give students the sense of randomness or unpredictability. The first part of the visualization was realized as an interactive web page with the simulation coded in JavaScript. Graphics were produced with Adobe Illustrator and Blender3D. The second part of the visualization is similar to the visualization in Paper III and shows the selection of resistant mutants in the form of colonies growing on two agar plates treated with antibiotics (Figure 13). Two control plates without antibiotics were also shown for reference. This part of the visualization was provided as still images for time restrictions reasons in development.

Del 1

- Ni har nu en enkel simulering framför er som visar hur ofta en viss mutation som ger ökad resistens mot antibiotika uppkommer i en gen hos en bakterieart. Bakterierna odlas i näringslösning i provrör UTAN antibiotika.
- Bakterierna i alla provrör kommer från ett enda prov med icke-resistenta bakterier.
- Om en mutation har skett när bakterierna har odlats i ett provrör markeras det provröret med en röd ring.
- En omgång av rör visas på skärmen när ni börjar. För att gå vidare om försöket klickar ni på "Ny omgång".
- Besvara nu frågorna på det utdelade pappret med hjälp av simuleringen.

Figure 12. Part 1 of Visualization 1 showing four runs of the bacteria cultivation experiment. Test tubes containing a mutation for antibiotic resistance are denoted with a red circle.

Figure 13. Part 2 of Visualization 1 showing the result after treatment of agar plates with antibiotics.

Visualization 2

Despite our initial analyses of the results from testing Visualization 1 were promising, the data also indicated that students had difficulties to understand that the presence of selection pressure (antibiotics) did not impact the mutation rate. We sought to address this by constructing another visualization that should clarify this. Therefore, we chose the context of deer mice and fur color evolution. The hypothesized advantage was that the selection pressure (light beach environment) would not affect the survival of the mice pups directly in the nest, compared to antibiotics which instantly affects a new bacterium after cell division. In addition, we hypothesized that it would be meaningful for students to compare the mutation rate in the presence and absence of selection pressure, in order to understanding the meaning of random mutations. As for Visualization 1, the simulations were implemented in JavaScript and the output and interaction components were presented as a regular HTML-web page. This also has advantages in terms of cross-platform compatibility and easy distribution, the application is also runnable on tablets and cell phones.

The visualization is comprised of three sections: 1) Random mutations of light fur colour in the absence of selection pressure 2) Comparison of random mutations in the presence and absence of selection pressure (forest and beach environment) and 3) Differential survival of dark and light mice in the two environments. Mutation frequencies were exaggerated in order to make the visualization usable. This was clearly stated in the on-screen instruction in order to avoid misconceptions. The survival probabilities of the dark and light phenotypes were based on field experiments (Hoekstra, 2011).

Below, a number of screenshots are provided that exemplify typical results of the simulation.

Figure 14. Introductory screen to Visualization 2. The section provides an overview of the topic and background information about deer mice and their ecology.

Figure 15. Comparison of mutation rates in two different environments (left panel: forest, right panel: beach).

Figure 16. Upper panel: Differential survival in the two different environments resulting in change in allele frequencies in the populations. Lower panel: Image shows the hiding function of dead mice which enables the user to easier visualize the resulting proportion of surviving phenotypes.

Ethical considerations

The research conducted in this thesis has mainly been financed by The Swedish Research Council (VR, Ref nr 2012:5344, LT) and the rules and guidelines applying for this research has been guided in the thesis and its related works. For research financed by VR, ethical guidelines apply. The major ethical guidelines followed for research in Sweden is the The European Code of Conduct for Research Integrity (Allea, 2017). The principles in the guidelines gives the researcher guidance on Reliability, Honesty, Respect and Accountability.

In addition, Swedish law applies to certain areas of research, for example data protection law. In the application for the EvoVis project, a plan for handling ethical issues was submitted and approved by the VR. The research conducted within the project has followed these guidelines. According to Swedish law, some research on human subjects requires permission from an ethical review committee. While the research in this thesis involves data collection from human subjects, the law is not applicable due to the fact that the research neither involves handling of sensitive personal data or physical experimentation on human subjects or biological material from a physical person.

For educational research there are ethical guidelines to be followed that safeguards the rights of the participants. A major requirement is “Informed Consent” in participation which means that participants should be duly informed about the purpose of the research, the types of data collected and how the material collected will be handled after the data collection. Also, participants must be informed that participation is voluntary and that they have right to withdraw from the study at any time without giving a reason (Robson, 2011). If a participant chose to withdraw, data connected to the individual should be deleted or returned to the participants. Data handling guidelines from Linköping University were followed to safeguard anonymity and data security while enabling tracking of data to an individual in the instance she or he chose to drop out. The guidelines follow the GDPR requirements of personal data handling.

Information and consent were handled by both oral and written information to the participants (see Appendix for example of consent form). In the case a participant was under 16 years of age, consent from parents was also collected. Information was sent to participants in advance of data collection mainly through their teachers. Before the data collections started, the researchers again informed about the data collection, data handling, voluntary participation and the right to withdraw at any time.

Results

This chapter presents an overview of results of studies underlying this thesis, which were conducted in a sequential fashion to first uncover the current status of visualizations for teaching and learning evolution before conducting design experiments. This process was guided by the Model of Educational Reconstruction (Duit et al., 2012) and follows components 2 and 3 of the framework. First, the results of studies relating to MER component 2 (teaching and learning) are summarized. The first phase explored the conceptual content of evolution visualizations (Paper I) and students' explanations of natural selection with an emphasis on threshold concepts (Paper II). The results of the first phase were synthesized and used as inputs for the second phase, in which novel visualizations were designed and evaluated (Papers III, IV and V).

Research on teaching and learning

The aim of this component was to explore learners' conceptions of the scientific content described in the theoretical background. In addition, visualizations (online videos) of natural selection were characterized in terms of their conceptual content.

The research questions guiding this phase were:

- How is natural selection *explained* in online videos? (Paper I)
- Do videos use the potential to explain the invisible or abstract concepts of evolution? (Paper I)
- How do students *explain* evolution in terms of key and threshold concepts? (Paper II)

These over-arching questions guided the first phase of the research, which was reported in two published papers (Paper I and II). Results of the studies presented in the papers are summarized in following sections.

Paper I

“A conceptual characterization of online videos explaining natural selection.”

Study I, reported in this paper, was a descriptive and mainly quantitative exploration of the degree that the communicated content of evolution visualizations match key and threshold concepts derived from the

scholarly literature. In addition, we analysed the presence of misconceptions and the organismal contexts of evolution in sampled visualizations (60 online videos). The following research questions were specifically addressed in this paper, in addition to the overarching questions:

1. What content with regard to specified key concepts, threshold concepts, and misconceptions associated with natural selection is conveyed by videos available on the internet purporting to explain evolution?
2. What organismal contexts are used to illustrate natural selection in videos available on the internet?
3. What patterns of co-occurrence of the identified content can be distinguished in the sampled videos?

Principles and key concepts of natural selection

The analysis revealed that the three principles sufficient for a basic explanation of natural selection were present in varying percentages of the videos: *variation* (72%), *selection* (60%) and *inheritance* (37%). However, the completeness of explanations in specific videos in terms of these three principles was lower. The most common combination of the principles was *variation* and *selection*, which were simultaneously present in over half of the videos. 20% of the videos did not include any of the principles and only a third (32%) included all three principles. In terms of the individual key concepts of natural selection, only one of the 60 videos (2%) included all concepts. Notably, only a third (32%) of the videos explained how novel variation occurs (*origin of variation*).

Threshold concepts

In terms of spatial scales, most of the videos represented the individual level (75%) followed by the population level (43%). However, in terms of the “imperceptible scales”, fewer videos dealt with these, particularly the genomic (30%) and molecular (13%) level. Similarly, temporal scales were mostly represented almost exclusively in years or generations. Concepts relating to the stochastic aspects of evolution were also only present in a minority of the videos: probability (30%) and randomness (23%). In addition, the latter two concepts were mainly communicated only orally (14%), and only 3,1% of the videos showed randomness visually, and one video of 127 (0.8%) presented randomness in both visual and oral modalities (unpublished data, see Appendix I. In addition, our preliminary analysis of the unpublished data showed that links between spatial scale levels were scarce (see Appendix I).

Representation of organismal contexts

A clear majority of the videos (70%) represented evolution of animals and 25% human evolution. In contrast, only 8% concerned plant evolution, 18% bacterial evolution and 10% used symbolic organisms.

Identified misconceptions

While the analyzed videos conveyed several evolutionary concepts, the analysis also revealed the presence of misconceptions. Most of the videos (nearly 80%), contained no misconceptions. The most frequent misconceptions found were *need-based explanations* (10%) and *inheritance of only beneficial traits* (8.3%) (i.e., selection leads to a genetically uniform population in one step).

Summary of results

How is natural selection explained in online videos?

The explanations in the analyzed videos can be characterized overall by incompleteness in terms of the basic principles of natural selection. Furthermore, their representation of threshold concepts is biased towards the perceptible scale levels (individuals and populations). Concepts related to stochastic aspects of evolution (randomness and probability) together with genetic and molecular level explanations were only present in a minority of the videos. Encouragingly, only a minority of the videos conveyed misconceptions about evolution.

Do videos use the potential to represent invisible or abstract concepts?

Our results from the analyzed videos show a strong bias towards the perceptible scales of space and time. Only a minority represent threshold concepts such as *randomness* or *probability*. Objects, events and processes on *spatial* and *temporal scales* outside perception are seldom represented (e.g., the molecular or genome scale levels and short or long temporal scales). The results therefore suggest that visualizations of evolution can be improved in terms of their representation of important conceptual content and that their potential is underexploited.

Paper II

“Conceptual characterization of threshold concepts in student explanations of evolution by natural selection and effects of item context.”

The aim of the study presented in Paper II was to explore how and to what extent students used threshold concepts in their explanations of natural selection. In addition, we explored whether use of threshold concepts was affected by item context in a similar manner to key concepts in previous studies. Three often-used test items on natural selection were used, situated in different evolutionary contexts (differing in organisms involved and/or trait gain/loss). Three research questions guided the study, in addition to the overarching questions:

1. How do students apply and express threshold concepts in their written explanations of evolution by natural selection?
2. How consistent is the use of threshold concepts across examples of natural selection with differing surface features?
3. Can relations between items' surface features and the students' expression of threshold concepts be discerned?

Threshold concepts

Generally, there were between-item differences in frequencies of threshold concepts used in the students' explanations, as expected. The most frequently used threshold concepts were spatial and temporal scales, while randomness and probability were more seldom used across the items (Figure 17). Almost a third (28%) of the students mentioned *randomness* in the antibiotic resistance context, significantly⁵ higher than in the other two contexts. The antibiotic resistance context also elicited the highest frequency of references to spatial scale.

In addition to the data on tertiary students published in Paper II, data was also collected and analysed for upper-secondary students from Sweden (unpublished data, see Figure 17). As expected, the frequencies of threshold concepts were substantially lower for the upper-secondary sample compared to university-level students. Most notable, *randomness* was almost absent in upper secondary students' explanations, while the other threshold concepts followed approximately a similar pattern to the tertiary sample across the items, although in lower frequencies.

⁵ Cochran's Q/Dunn's post hoc test with Bonferroni correction for number of comparisons. See Paper II Methods section for further details.

Figure 17. Use of threshold concepts use in explanations of natural selection in response to the 'Bacteria', 'Cheetah' and 'Salamander' items by upper-secondary students (upper panel, unpublished data, n=75 students) and tertiary students (lower panel, Paper II, n=247 students).

Key concepts

Generally, use of key concepts was also affected by item's context (See Figure 3 in Paper II). *Origin of variation* followed the pattern of randomness and spatial scale, being used significantly more frequently in responses to the Bacteria item than in responses to the Cheetah and Salamander items. In contrast, presence of variation within a population (*individual variation*), as well as *differential survival* and *reproductive success*, were significantly more common in responses to the Cheetah item than in responses to the other items. Interestingly, *inheritance of traits* and *change in population* were least frequently present in students' explanations.

Consistency across items

The consistency of concept application across all three items was generally moderate (see Table 3 in Paper II). The most consistently applied concepts across all items were selection pressure and individual variation (respectively used by 44 and 26% of students in responses to all items). Notably, concepts linked to appearance of novel variation on the genetic level — origin of variation, randomness and spatial scale — were very seldomly used across all items (by just 13, 2 and 5% of the students, respectively). Similarly, threshold concepts were less consistently applied across all items than key concepts (see Table 3 in Paper II,). Moreover, the concept that random mutations are a source of novel variation was used at most by 20% of the students (in responses to the Bacteria item, see Table 4 in Paper II).

Summary of key results

Generally, frequencies of use of both key and threshold concepts in students' explanations of natural selection were context-dependent. To my knowledge, Paper II provides the first report of contextual effects on use of threshold concepts. Overall, threshold concepts were used less frequently than key concepts in the explanations. A notable issue with the explanations, especially those of upper-secondary students, is the scarce use of randomness. Moreover, the item with a microbial (bacterial) context elicited most use of spatial scale and randomness in explanations while the Cheetah item yielded most mentions of probability. Use of individual variation (most common in the Cheetah item) did not follow the same pattern as origin of variation (most common in the Bacteria item). Not surprisingly, *differential survival* and *reproductive success* followed

the same pattern as individual variation (although at lower frequencies), being most common in the Cheetah item. Threshold concepts generally occurred rarely in learners’ explanations and less frequently than the commonly analysed key concepts

Synthesis of results presented in Papers I and II

The results presented in these papers enable some comparisons of the representation of key and threshold concepts in visualizations and students’ explanations. In Figure 18, frequencies of key and threshold concepts in online videos and student explanations are compared. Although this comparison has limitations (primarily that students’ explanations are based on only three examples while the videos cover diverse examples of evolution), some general patterns can be discerned. First, stochastic concepts (randomness and probability) are more rarely present in both online videos and student explanations than common key concepts. Second, genome-level phenomena are seldom used in either videos or student explanations. Accordingly, key concepts related to genetics (*origin of variation* and *inheritance*) were less frequent than *individual variation*, *differential fitness* and *selection pressure* in both videos and students’ explanations (see Figure 18).

Figure 18. Comparison of concept frequencies from Paper I (Online videos about evolution, n=60 videos) and Paper II (Students explanations of natural selection, n=247 university students). (Note that due to slightly different analytical categories between the papers, selection pressure and differential fitness variables are pooled.)

Implications for design of learning environments

The results presented in Paper I clearly showed that visualizations seldom represent the concepts that appeared to be most challenging for students in Study II (*randomness, probability, spatial scale, molecular and genome levels*). Although comparison of the data acquired in Studies I and II has certain limitations and the relation is purely associative, this indicates a potential focus for improving the design of visualizations for learning evolution. The findings that students tend to construe evolution as a goal-directed or need-driven process and have difficulties in discerning the role of origin of variation and natural selection suggested that addressing the role of random mutations may be crucial. Thus, the development of novel visualizations in the second phase of studies underlying this thesis (reported in Papers III, IV and V) focused on exploration of visualization design to convey the roles of randomness and genome-level phenomena in evolution.

Design and evaluation of learning environments

According to the MER framework, science content structure has to be transformed into a content structure for instruction. Papers I and II confirmed the hypothesis that certain concepts that are important for understanding evolution (randomness, probability and spatial scales) were not emphasized in available visualizations and students seldom use these concepts to explain natural selection. Hence, Papers I and II indicated that a content structure for instruction should include and emphasize these concepts. Therefore, objectives in the second phase of the empirical investigation were to design visualizations that foster students' learning and understanding of these concepts and evaluate the designs.

For designing learning environments such as visualizations, the MER framework emphasizes the importance of making the content accessible to students. The key idea emerging from Studies I and II was that students should benefit from visualizations stressing the role of random genetic variation in natural selection. The framework also emphasizes that the content for instruction should be enriched by putting the ideas into a context that makes sense or is relatable for the students (R. Duit et al., 2012). In the developed visualizations this was done by choosing antibiotic resistance: a context that is familiar to students because most (if not all) students have personal experience of taking antibiotics and some understanding of their importance for society. In addition, antibiotic resistance poses mounting threats that many students have heard about. Thus, as a societal issue of vital importance for humans in

the future, we hoped that students would find it interesting and meaningful to learn more about the topic.

MER is a recursive process in which the components do not necessarily proceed in a sequential fashion, and in a MER-based experimental teaching study the outcome informs revisions of the learning material, or in this case visualization design. As described in the following sections, this feedback mechanism was applied in Studies III-V, in which three visualizations were successively developed and evaluations of the first and second fed into the design of the next visualization.

Paper III

“Insights from introducing natural selection to novices using animations of antibiotic resistance.”

The aim of the study presented in Paper III was to develop and pilot an interactive visualization of the evolution of antibiotic resistance. It was mainly an experimental and qualitative study due to its explorative nature. An interactive visualization was developed (see the Methodology chapter for a description of the design and development) and tested with 32 secondary grade 8 school students (13-14 years). The following specific research questions were addressed in the study:

1. What characterizes novice pupils’ understanding of the origin of resistance, and how is their understanding affected by an interactive animation and accompanying exercise?
2. What obstacles and/or opportunities can be discerned for teaching the evolution of antibiotic resistance to novice pupils through a series of interactive animations in terms of (a) the origin of resistance, and (b) the ability to make predictions?
3. What evolutionary aspects do the pupils choose to include when asked to transfer reasoning from a bacterial to a mammalian context?

Results

Most of the students successfully provided evolutionary explanations in the post-test, but not pre-test (Figure 19). The difference (an increase in frequency from 6 to 17 of 32) was very significant according to the McNemar-test ($p < 0.007$). Predictions about evolution of resistance: In addition, seven of the eight groups formulated reasonable predictions about the evolution of resistance, i.e., they correctly predicted that not all bacteria in the experiment would be killed by antibiotics and some

colonies would appear, although fewer than on control (antibiotic-free) plates (see Table 3 in Paper III).

Two major themes explaining origin of variation were detected in transcripts of the groups' interactions: variation by mutations and teleological explanations. Most groups (six of eight) explained novel variation as arising from mutations while two attributed resistance to an active (teleological) choice of the bacteria. Four groups also mentioned that mutations occur in DNA during replication and cell division. In addition, six groups acknowledged that variation is heritable by cell division and that this is a rapid process in bacteria. Mutations were mentioned as errors or nucleotide mismatches. According to the teleological explanations, bacteria learn how to deal with antibiotics, get used to them, or they want to live and reproduce (implying that bacteria have goals). In addition, the depicted nucleotides in the visualization of DNA replication were interpreted as agents, or programmed, or things that knew where to be placed.

Figure 19. Distribution of responses to the pre-post item used in Paper III. The correct option (C) was selected by a significantly higher proportion of students in the post-test according to the McNemar test ($p < 0.007$). (For A,B,D – refer to Methodology section for a description)

Influence of the animation on students' reasoning

In several instances students directly referred to the visualization, especially the section depicting DNA-related events. For example, two groups referred to the base-pairing errors. One student mentioned that visualization helped their understanding because it demonstrated randomness and four justified their change in response to the post-test item with the

visualization. However, it should also be noted that one student chose not to change her response in the post-test because she did not understand the visualization.

Transfer to a mammalian example

In order to explore possible transfer of any learning of evolutionary principles, an open-ended post-item was used, asking students to compare bacterial evolution of antibiotic resistance and giraffes' long necks. While none of the students included variation, selection and inheritance in their responses, 16 of 32 mentioned changes in DNA or genes in their responses. Eight students explicitly mentioned common features of giraffes and bacteria, often DNA or genetic changes. In addition, four explicitly linked changes on the genetic level and phenotypic changes, and one provided a fairly complete response mentioning *randomness*, *genetic changing leading to phenotypic change (origin of variation, spatial scale)*, *inheritance* and *change in population*. It should also be noted that four students based their answers on teleological thinking, e.g., that bacteria and giraffes evolve because of a need to survive.

Summary of results

Overall, the results indicate that the visualization had a positive effect on students' understanding and reasoning about evolution, in terms of both individual post-test results and responses to the group questions. This is encouraging since previous studies have shown that antibiotic resistance is difficult to explain for students of the same age in national tests at the end of the biology course for compulsory school (Lind Pantzare & Abrahamsson, 2015). The fact that some students were able to transfer knowledge gained from the visualization to a mammalian context is also a positive outcome, especially since transfer of knowledge of natural selection to other contexts is often difficult for students (Nehm & Ha, 2011). The focus on the role of genetic variation presented in the visualization appears to have been fruitful for transfer. 50% of the students mentioned DNA or genetic changes as a common denominator between the bacterial and giraffe contexts. Understanding that evolution applies to all organisms because they share the same fundamental genetic machinery could thus facilitate transfer between contexts. Despite the successful outcome, the presence of teleological reasoning in some responses shows that a short intervention of this type does not necessarily lead to all students changing their explanations of natural selection.

The tendency for some students to apply teleological explanations about molecules is intriguing but has also been found in other studies

(Talanquer, 2007). Our hypothesis was that showing molecular level events as a cause for variation would prompt students to abandon teleological explanations. As there was no control group in the study, it cannot be definitively established whether the visualization affected this tendency, but a substantial percentage of the students did not use teleological explanations following exposure to it. Interestingly, transcript data showed that despite seeing depictions of random events such as the Brownian motion of nucleotides, students may still ascribe agency or teleological explanations even to molecules. Whether this was metaphoric or literal cannot be concluded from this study, but should be further explored in future studies. In addition, the visualization did not depict the concept that random mutations are neither caused nor affected by selection pressure, a common misconception. Comparisons of mutation frequencies in the presence and absence of selection pressure might help students to understand the intended sense of random mutation better.

Paper IV

“Development and evaluation of digital interactive simulations of randomness in evolution.”

Design considerations

While the visualization in Paper III showed random movements of nucleotides, the students who participated did not seem to infer that mutations must be random. Thus, in Paper IV, we decided to further explore means of visually representing randomness of mutations. We were also interested in how students’ intuitive interpretations of antibiotic resistance interacted with the presented scientific content. Since an intentionally triggered process can have random outcomes (cf. a dice throw in a board game), we also decided to stress a key role of randomness in evolutionary contexts: mutations’ independence of selection pressure. The design hypothesis was that randomizing the appearance of novel mutations in the visualization would make students use this threshold concept more frequently in the post-test. To further strengthen the impression of randomness, the visualization was designed to enable the user to create new batches of test tubes with bacteria and observe the occurrence of mutations in different tubes and numbers in each batch.

To highlight the concept that mutations are not directed, i.e., not a triggered response to selective pressure, the students were told that the test tubes did not contain any antibiotics. In the second part of the visualization, the aim was to show how antibiotics act as a selective factor and reinforce understanding that mutations can also appear in the absence of antibiotics. The visualization shows four agar plates with

bacteria sampled from the test tubes in the first part of the visualization (two with and two without supplementary antibiotics).

Pre- and post-test results

Results of the pre- and post-test used are presented in Figure 20 . Following exposure to Visualization 1 we saw a large increase in the number of students using *randomness* in their explanations of natural selection after the intervention, while Visualization 2 showed virtually no difference between the pre- and post-tests.

Figure 20. Threshold and key concept use on the pre- and posttests in Visualization 1 (Antibiotic resistance) and Visualization 2 (Deer mice fur color).

In the pre-tests, no students explained the origin of a novel trait with random events. In contrast, a substantial percentage (33%) used the concept *randomness* in their responses in the post-test, but not following

exposure to Visualization 2. This increase was paralleled by an increase in mentions of genetic explanations connected to *origin of variation* (from 30% to 63%).

Results from group tasks

The results from the group tasks showed that most groups inferred that mutations occurred at random in the visualizations. When asked about whether the presence of selection pressure (antibiotics in Visualization 1) would affect the frequency of adaptive mutations, several groups correctly stated that this would not alter the probability of mutations. Some groups however suggested that the rate of adaptive mutations would increase if selection pressure was present. Following exposure to visualization 2, where a special section focused on comparison of adaptive mutation rates in the presence and absence of selection pressure, most groups predicted that selection pressure would increase the rate of adaptive mutations. Accordingly, several groups were surprised by the results showing no difference in adaptive mutation rate due to selection pressure. In addition, several groups concluded that they could not explain why this was the case.

Explanations of the origin of variation

Three types of explanations for the presence of variation could be discerned during students' interactions as they engaged in group tasks associated with Visualization 1: *Mutations cause variation*, *Variation occurs due to selection pressure* (exposure to antibiotics), and *Existing variation*. However, after completing the second part of Visualization 1 (showing selection and indicating the presence of mutations prior to antibiotic exposure) all but one of the groups concluded that antibiotic exposure was not necessary for resistance-conferring mutations to occur. Some of the groups referred to the visualization as a source for their argument. The finding that many groups expected the presence of selection pressure to influence the rate of adaptive mutations is an indication of their construal of mutations as directed or affected by the environment.

Conceptual change in explanations

In addition to analyzing the use of isolated concepts in the pre- and post-tests, we categorized students' models according to their degree of conceptual change: *No model*, *Intuitive model*, *Synthetic model*, *Mixed model* and *Scientific model*. Synthetic and mixed models combine elements of intuitive or everyday explanations with scientific ideas, e.g.,

need-based adaptation mediated by a genetic mutation. Results of the categorization is presented in Figure 21. A *synthetic model* was most common in responses in the pre-test, but a *scientific model* was the most common in the post-test responses, following exposure to Visualization 1. In contrast, there was a less marked increase in use of a scientific model in the post-test following exposure to Visualization 2, although frequencies of *no model*, *intuitive model* and *mixed model* decreased. Here, a *synthetic model* was the most frequent in both pre- and post-test responses.

Figure 21. Categories of conceptual models detected in students' responses in pre- and post-tests following exposure to Visualizations 1 (left panel) and 2 (right panel).

Causal models used by students

In-depth analysis of the causal models was mainly aimed to uncover the nature of students' synthetic and intuitive models. In the pre-tests, two major types of explanations dominated prior to Visualization 1 (see Figure 22 and Paper IV, Table 2): *Environmentally induced changed in traits* (33%) and *Scientific explanations* (23%). Immunization was commonly used as an analogy or mechanism in the first category. For the post-test transfer item (Cheetah) the pattern of causal models was different with scientific explanation dominating responses (57%) followed by need-based reasoning (23%).

Need-based (43%) and *Gene-based without a cause* (26%) models dominated in responses before exposure to Visualization 2. In addition, no scientific explanations were found in responses to the pre-test. Although the frequency of scientific explanations increased in the post-test (to 17%), need-based reasoning only decreased to a minor degree. Surprisingly mutations were also rarely mentioned after the visualization.

Figure 22. Major categories of students' alternative explanations about natural selection: A) Need-based adaptation, B) Need-based adaptation mediated by mutations (synthetic model), C) Need-based adaptation combined with differential survival (synthetic model).

In conclusion, the visualization of randomness appears promising. Results of the group tasks indicated that most students perceived randomness in both visualizations. However, there was a difference between the two visualizations in frequencies of students who mentioned randomness in their post-test explanations. This difference is not necessarily an effect of the different visualizations and may be due to differences in pre-knowledge of the participating students, and/or characteristics of the pre- and post-tests.

Paper V

“Framing evolution – the interplay between students’ intuitive conceptions and scientific explanations of evolution.”

Paper V builds on the same data set as Paper IV (Dataset IV). In this paper, the organizing semantic frames of student explanations were identified. Conceptual blending theory was used to analyze how students integrated scientific elements with the semantic frames of their explanations. The context of antibiotic resistance activates a certain repertoire of pre-knowledge (context-dependence) such as immunization, vaccination, learning or tolerance. These frames continued to be used by many students after the intervention and scientific elements presented in the visualization were organized into the frame. Thus, the conceptualization before using the visualization affected the construction of the explanation and indicates that semantic frames underlie synthetic explanations. The explanations used by many students after the intervention could be characterized as synthetic conceptions. In some instances, the student explanations were identified as mixed in the analysis in Paper IV. These explanations used two different frames and it was apparent that the pre-conception was not compatible with the information that mutations conferring antibiotic resistance occurred before the exposure. This was solved by some students by stating two different explanations in their responses: that resistance either was due to exposure OR due to mutations occurring before exposure to antibiotics. These two explanations are clearly organized by two different frames with different causal order of the events. Thus, mixed explanations might occur when students are not able to create a synthetic explanation by organizing all the elements of the taught explanation into the same semantic frame.

Discussion

This discussion is based on the results previously presented in the papers. The research questions initially posed in the thesis are discussed in relation to the results and existing literature. The focus of the discussion is the representation and understanding of randomness, which turned out to be the major obstacles for students.

Representation of natural selection in visual media

The results in Paper I showed that the basic principles of natural selection (variation, selection and inheritance) were present in a large percentage of the analyzed videos. However, the principles were seldom represented in the same video. In terms of key concepts, only a fraction of the videos represented all key concepts of natural selection.

Several videos have the potential to communicate natural selection but not the origin of variation. Since there exists no other studies of this kind, comparisons of our results are not possible. However, surveys of other educational media for evolution teaching such as textbooks have revealed that misconceptions and errors are present in evolution textbooks (Aleixandre, 1994; Tshuma & Sanders, 2015) or that information is segregated in different chapters (Nehm et al., 2009). Catley & Novick (2008) surveyed images of evolutionary diagrams in undergraduate textbooks, finding that cladograms were dominating. In addition, many diagrams were deemed inaccurate or judged to create interpretation difficulties for students. This puts additional demands on teachers in selecting and critically examining textbook content for instruction.

The fact that both visualizations and textbooks tend to be incomplete in terms of basic evolutionary concepts and that concepts are not presented in relation to each other is likely to inhibit the development of correct mental models of natural selection. These shortcomings also means that teachers must not rely heavily on learning media but put additional effort and attention to guide students' development of correct mental models.

Another issue with the analyzed videos was the biased representation of biological taxa. It is known that students have difficulties to reason about evolution in different contexts (Ha & Nehm, 2013; Heredia et al., 2016; Ross H. Nehm & Ha, 2011). Thus, teachers should carefully select videos as well as other instructional examples of natural selection and actively help students to compare different cases.

To what extent do visual media use the potential to represent abstract concepts of evolution such as randomness, probability spatial and temporal scales?

In addition to the analysis of key concepts in videos, the analysis in Paper I revealed that a majority of the videos fail to use the potential of dynamic visual media to represent threshold concepts. Thus, their representational potential is underutilized. Regarding randomness and origin of variation, the perhaps most challenging concepts for students, these were seldom represented in the videos. Our additional analysis of the mode (visual or oral) of concept representation (unpublished data, see Appendix X) revealed that a mere 3% of the videos represented randomness in a visual and thus concrete manner. Given the difficulties for students to understand randomness in evolution (Garvin-Doxas & Klymkowsky, 2008), this is troubling. In addition, previous research indicates that evolution textbooks seldom address probabilistic reasoning and chance (Aleixandre, 1994).

The potential of visualizations to represent objects and processes outside our perceptual ranges in time and space was sparsely used in the analysed videos. For example, mutations or other events on the molecular scale were only present in 13% of the videos. Considering the definition of natural selection as change in alleles frequencies in a population, it was surprising to find that the genome level was only mentioned in less than a third of the videos. Instead, the macro level dominated with individual and population levels. In addition, links between different spatial scale levels were seldom explicated. Making causal connections between and translating across levels of biological organization is considered to be central for students abilities to explain biological phenomena (Jördens et al., 2016; Tsui & Treagust, 2013) but students often have difficulties (e.g. Duncan & Reiser, 2007; Ferrari & Chi, 1998; Marbach-Ad & Stavy, 2000; Shtulman, 2006). Hence, our results indicate several areas for improvement of evolution visualizations.

Threshold concepts in students' explanations

Numerous research papers have assessed students understanding of natural selection in terms of key concepts (see Kuschmierz et al., (2020) for an extensive review but also Bishop & Anderson (1990); Federer et al. (2015); Nehm & Ha (2011); Nehm & Reilly (2007)). As previously argued, these key concepts fail to account for important aspects of evolution such as randomness, probability, spatial and temporal scales, so called threshold concepts. Not understanding and relating these concepts to natural selection in a correct way can lead to misconceptions. However, to date no structured survey exists of how students use these concepts in their

explanations of natural selection. Paper II represent a first study in this direction. In Paper II, the following research question was addressed:

How do students use threshold concepts in written explanations of natural selection?

The results shows that university students generally user threshold concepts less frequently than the commonly analyzed key concepts, with randomness and probability being the least used. Randomness, probability and spatial scale also showed to be tied to item context. This indicates that well-known item context effect for key concepts (Federer et al., 2015; Heredia et al., 2016; Nehm & Ha, 2011), also applies to threshold concepts. A general issue with transfer of natural selection between contexts for students is to realize that DNA and genes are common denominators across all living taxa. Hence, instructional strategies focusing on comparing examples of evolution in a diverse of biological taxa with a DNA-focus should be beneficial and has previously been advocated (Jördens et al., 2016). An interesting result on spatial scale and randomness is the significantly higher presence of these concepts in explanations of antibiotic resistance than for the other two surveyed items (cheetah running speed and loss of sight in cave salamanders). A tentative explanation is that there are fewer organizational levels to link in the bacteria example (cell level and individual levels are identical). Another possibility is that students are less likely to ascribe intentional actions to micro-organisms than for animals who possess high-level cognitive functions. If so, micro-organism context can have affordances as an introductory context of evolution.

The effect of item context on probability with the cheetah item have the highest frequency was largely explained by the reasoning about survival probability and chances of catching preys. Thus, probability occurred most frequently in connection with differential survival and in the cheetah item. Probability was also found in connection with other concepts (reproductive success, origin of variation and inheritance) but to only to a minor degree. This indicates that students do not reason about natural selection as a probabilistic process in general and agrees with previous findings where natural selection is described in deterministic manner by both students (Gregory, 2009; Mead & Scott, 2010; Nehm & Reilly, 2007) and in textbooks (Aleixandre, 1994). Thus, our findings lend further weight to the importance of guiding students towards a more stochastic and probabilistic conception of natural selection.

The role of interactive visualizations

How do interactive visualizations focusing on threshold concepts, especially randomness and mutations, affect students understanding of evolution?

The results imply that showing randomness in a concrete manner (although having a positive effect on post-test explanations in Visualization 1) does not necessarily mean that students experience a conceptual conflict. Random process can still be under the influence of an agent. Compare for example the every-day experience of rolling a dice in a game. The player clearly has an intention of achieving a goal and actively triggers the dice roll. Although the outcome of the dice roll is random, rolling the dice many times will sooner or later lead to the desired outcome. A known problem with randomness is that people easily infer mechanisms, causal connections or patterns (Rominger et al., 2011). Conceptions or views about nature and the biological world can also affect the interpretation (Evans, 2001), thus making it difficult for learners to change their understanding. Guiding and reflection should be important as well as the statistical reasoning ability. For example, repeating an experiment a number of times. The working in groups appears to facilitate by comparisons of results and also discussion and reflection.

Comparisons between Visualization 1 and 2 must be done with some caution due to a number of limitations. First, the student samples differ in education level and age by one year (grade 9 vs grade 10). Second, the context of the test items are known to affect the pattern of key concepts (Göransson et al., 2020; Nehm & Ha, 2011), threshold concepts (Göransson et al., 2020) and misconceptions (Nehm & Ha, 2011). Accordingly, we used the same item for both the pre- and post-test. For Visualization 1, the pre- and post-test item used was directly connected to the context of the visualization (antibiotic resistance) and for Visualization 2 the cheetah item was used for the pre- and post-test. Thus, for Visualization 2, the pre- and post-test item did not ask about an identical context of evolution (change in running speed vs change in fur colour). In visualization 2, a second item was used on the post-test that related to the context of the visualization, evolution from brown bear to polar bear. This item asked about a similar change in trait (fur colour), however the species were different, and the item context was between species change compared to intra-species change in the visualization. Thus, the contexts of the items and visualizations differ between Visualization 1 and 2. Therefore, absolute comparison of concept frequencies should be done with caution. However, pre- to post changes can to some extent be

compared, bearing in mind that the choice of items in Visualization likely underestimates the degree of learning.

The conceptual conflict likely experienced between mutations as random and preconceptions of evolution as a directed process appears to be resolved by some students by constructing synthetic explanations. Some students construct synthetic explanations where mutations become the mechanism mediating the change, but the change is triggered or caused by the environment or need. Yet others use mixed explanations where the taught idea of random or non-directed mutations is presented as an alternative explanation to the intuitive idea of directed mutations or adaptation.

Interaction between preknowledge and taught content

How are students' explanations of natural selection related to semantic frames and conceptual blends?

It is well-known that students often fail to explain evolution by natural selection, such as antibiotic resistance, despite carefully designed teaching (e.g., Bishop & Anderson, 1990; Demastes et al., 1995; Govindan, 2018; Gregory, 2009; Nehm et al., 2010; Nehm & Reilly, 2007). Often students tend to harbor misconceptions, and these have been suggested to hamper their learning of evolutionary principles (Barnes et al., 2017; Trommler & Hammann, 2020). Misconceptions and naïve or intuitive ideas have been found to co-occur with scientific ideas or key concepts after instruction (Nehm & Ha, 2011; Nehm & Reilly, 2007). In this paper, we have made a first exploration using frame semantics to analyze students' conceptualization of antibiotic resistance before and after a teaching intervention. The results suggest that students often frame the evolution of antibiotic resistance as an adjusting or immunization process, in line with previous findings by other authors (Brookes-Howell et al., 2012; Govindan, 2018; Peel et al., 2019). In addition, the results also suggest that conceptual blending can describe how and why hybrid or synthetic conceptions form as the result of teaching. Such conceptions have previously been described in Vosniadou's framework theory of conceptual change (Vosniadou & Ioannides, 1998), but we believe that frame semantics and conceptual blending offers a valid and well-established cognitive mechanism that explains how and why synthetic conceptions are formed. In addition, conceptual blending can potentially account for some of the previously observed simultaneous presence of "naïve" ideas and scientific elements in student's natural selection explanations (Nehm & Ha, 2011; Nehm & Reilly, 2007). Another observed effect in evolution assessment is context effects on the number of key concepts and intuitive ideas used by students (Nehm & Ha, 2011). It is possible

that some of these effects are due to the different activation of frames by surface features of the context, and therefore we suggest that frame semantics could be used to explore item context effects of explanations.

Taken together, the results from Paper V indicates that activation of certain semantic frames are related to how students conceptualize taught content (see Figure 23). The frames appear to guide which elements of the scientific explanation are used to construct an understanding of the content. If the initially selected frame is not changed (c.f. frame-shifting by Coulson, (2001)), certain elements of the taught content may be disregarded or conceptualized in another way than intended. In Vosniadou’s terminology this is equal to a framework theory guiding the interpretation of content (Vosniadou, 2019). Thus, semantic frames and conceptual blending offers a cognitive linguistics approach to conceptual change which means that frames are the central organizing idea in meaning-making. As a consequence, the interplay between framing and for example visually representations (e.g., Powell et al., 2015) should be central for constructing visualizations for conceptual change.

It is our hope that this first study shows how frame semantics and conceptual blending can be valuable tools to enrichen our analytical toolbox for students’ conceptions and to describe conceptual change in terms of established cognitive theories.

Figure 23 Suggested model of framing and the construction of meaning from taught content.

Conclusions

A major finding from the studies in this thesis is the importance of considering the role of threshold concepts for students understanding of natural selection. The results show that students often fail to explain origin of variation by random mutations. In addition, students have difficulties in discerning origin of variation and natural selection as two separate processes. The relatively low presence of threshold concepts found in extant visual media explaining natural selection is troubling, given the aforementioned difficulties for students. Further troubling is the low number of visual media that have adequate quality for teaching in terms of complete representation of important principles and concepts for natural selection. In addition, the visual media analyzed as part of this thesis often fail to use the potential advantages of visualizations to represent abstract content of natural selection: objects and processes outside human perceptual ranges (e.g., molecules, genes or different time scales).

The development and evaluation of novel interactive visualizations based on the proposed content structure of key and threshold concepts shows that visualizations can indeed guide students towards a more scientific understanding of natural selection with regard to the role of random mutations. However, the results from the design experiment also indicates that even explicit instruction about the role of randomness is not sufficient in all cases for achieving conceptual change about natural selection.

The data analysed indicate that randomness may be an especially problematic concept to represent in an unambiguous manner. Top-down processes in visual interpretation may bias students to infer patterns in randomness which infers with a correct interpretation. Here, the results indicate that guiding students towards making structured observations with the help of e.g., guiding instructions, allowing them to predict the outcome and reflect and discuss upon it may be beneficial. Also, the results from the visual design studies indicate that students' preconceptions may lead to the construction of mixed or synthetic explanations of natural selection considering random mutations. Several examples in the analysed data shows that rather than changing conception from the often found need-based ideas to a scientific idea when taught about mutations, mutations were incorporated into synthetic models of natural selection where mutations were used as the mediating mechanism between selection pressure, need or environmental factors and heritable traits. This is at odds with the scientific explanation where mutations occur randomly but confirms previous findings where directed adaptive mutations have been reported.

The application of cognitive linguistics for analysing students' explanations offers a fruitful theoretical approach to understand why and how synthetic conceptions form. In addition, the approach also suggests that instructional design regarding language use and visualization design can benefit from considering framing theory and conceptual blending theory in the future.

The use of combination of quantitative and qualitative methods as well as deductive and inductive analyses was advantageous for revealing student conceptions in more detail. The often-favoured approach to evolution assessment by analysing the presence of isolated concepts should therefore benefit from being complemented with analyses capturing the causal reasoning patterns and the relation between different concepts in students' explanations.

The educational reconstruction of science content for instruction, including both key and threshold concepts has been productive both for research into students' conceptions but also for the design and evaluation of novel learning environments. The content structure may be of interest for curriculum designers, teachers, textbook authors but also for designers of educational media.

Implications and outlook

Implications for the design of visual media

The results in the thesis have implications for the design of visual media for instruction regarding the content and the representation of the content (e.g., visual and interaction design). In terms of conceptual content, our results from Paper I shows that designers should consider the basic principles of natural selection as well as origin of variation more carefully in their designs. Considering the representation of the content, the results show that representing randomness in origin of variation as in Paper IV is beneficial for guiding students' conceptual development. The visual comparison of mutation frequency in presence and absence of selection pressure in Paper IV also appear to help students to understand the meaning of randomness. This visual design could also be applied in other contexts than evolution where randomness is an important concept.

In addition, the short focus on natural selection (over one generation) in the visualizations should be expanded since many students did not improve on this from pre to post. Also, students framing of the visual content on randomness with regard to their preconceptions appear to bias them towards inferring patterns in randomness. Means to minimize such biases should be further explored in future designs, for example by revised instructions and reflection questions.

Implications for teaching evolution

A major finding in the results in this thesis is the difficulties students have with understanding the role of randomness in natural selection. Studies on teachers' pedagogical content knowledge imply that origin of variation has received relatively little attention compared to natural selection (Ziadie & Andrews, 2018). While this thesis is not concerned mainly with teachers' knowledge or teaching practices, some implications for teaching can be deduced from the results in terms of content selection for instruction assessment of students' knowledge of concepts related to variation and the selection and use of media for instruction.

Learning objectives and content in instruction

The results have some implications for learning goals for evolution instruction for secondary and upper-secondary education. The results suggest that origin of variation, although important, might be challenging

for students to understand. In addition, random mutations are also important in order to understand other evolutionary processes such as genetic drift. Evolution instruction should aim to convey at least the basic principles of natural selection (variation, selection and inheritance) together with origin of variation and randomness. Bearing the frequent synthetic conceptions seen in our and others results in mind that construe evolution as a directed causal process, it is important for teaching to stress the two-step process of evolution by natural selection: origin of variation and selection. The difficulties for students to integrate randomness in explanations of natural selection also indicates that random origin of variation and genes as a concept should be emphasized, so that students will have multiple opportunities to work with this content. In principle, there is no obstacle to introduce randomness in biology instruction already in primary school, for example by concrete means such as dices.

Assessment of student understanding

The results of the thesis also add to the extant body on assessment on student thinking about evolution. Paper II revealed that although students use a fair number of key concepts to explain natural selection, they may lack important threshold concept. Thus, it is important to assess students' knowledge of these concepts in addition to key principles and concepts. This is especially apparent for random mutations. Also, the results show that assessing threshold concepts with open items are subject to effects of item context. Teacher should consider using multiple items to assess students understanding of threshold concepts more adequately, due to the effects of item contexts revealed in Paper II.

Using media in natural selection instruction

The results from our studies on visualizations to communicate evolution and especially threshold concepts such as randomness and spatial scale indicates a number of issues that teachers should consider.

Regarding selection of visual media for instructional use, our results indicate that teachers should be careful when selecting visualizations from online sources, even when credible sources are used. Many visualizations are incomplete in their explanations and teachers should consider how the visualizations fit in the teaching sequence, what pre-knowledge students should have to meaningfully engage with the visual media. Furthermore, teachers should be aware that very few visual media cover the topic of natural selection completely in terms of important concepts. Teachers could consider using multiple visualization and/or complementing visuals with guidance and support.

The results also suggest that discussion questions guiding students' observations when engaging with visualizations are important. For example, when using simulations that include randomness it is important to instruct students to run simulation repeatedly and reflect upon that randomness means that the outcome will look differently each time. This may help students to avoid incorrect inferences guided by their preconceptions.

Future research

The results of the thesis also indicate potential areas for future research. While Paper I analyzed the conceptual content of online videos, we did not analyze what specific representations were used for each concept. Future studies using semiotic theory could analyze how natural selection concepts are represented in visualizations and how they are linked into explanations by means of narratives. Such research correlated with studies of students understanding of the content may reveal important results for future visualization design.

While Paper IV studies students' explanations of natural selection in relation to interactive visualizations of randomness, future studies should benefit from in-depth studies of students while working with the visualizations. Such studies should benefit from using e.g., video recordings and group discussions.

The importance of supporting material complementing visualizations was raised in Paper IV. Future studies should aim to refine supporting material that are effective in guiding students towards understanding the role of randomness in natural selection. Also, visualizations exploring multiple representations and zooming for understanding spatial scale links are interesting in order to help students link origin of variation on the genetic scale to individuals and populations. The often-cited problem of transfer of natural selection between contexts (and as revealed in this thesis) may be addressed by visualizations explicitly comparing natural selection in several taxa and traits. Here, the theory of multiple representations using dynamic linking is interesting to explore, as well as the significance of understanding the threshold concept of spatial scale (i.e., variation is caused by genetic variation across all taxa).

Lastly, studies of teachers PCK of threshold concepts in natural selection may also be informative for improving practice and teacher education.

Presentations

List of conference presentations made during the thesis work.

- Bohlin, G., 1981-, Author, Göransson, A. C., Author, Tibell, L. A. E., Author, Linköpings universitet, I. för teknik och naturvetenskap, Medie- och Informationsteknik, Originator, & Linköpings universitet, T. högskolan, Originator. (2014). *Evolution on the set – A conceptual characterization of online dynamic visualizations*.
- Göransson, A., Orraryd, D., & Tibell, L. (2015). *Searching for threshold concepts in evolution by using an open response instrument*. ESERA 2015 11th biannual Conference of the European Science Education Research Association. 31 august - 4 september, 2015, Helsinki, Finland.
- Göransson, A., Stenlund, J., & Tibell, L. (2014). An interactive visualization for communicating troublesome concepts of natural selection. *FND 2014, Forskning i Naturvetenskapernas Didaktik-Modeller Och Representationer i Naturvetenskaplig Forskning*, 5-6 November, Karlstad.
- Göransson, A., Stenlund, J., & Tibell, L. (2015). A Novel Computer Application for Teaching Evolutionary Mechanisms: Visual Analogies of Randomness and Natural Selection. *NARST 2015, Annual International Conference*. 11-14 April, 2015, Chicago, USA.

References

- Aldridge, M., & Dingwall, R. (2003). Teleology on Television? *European Journal of Communication*, 18(4), 435–453.
- Alexandre, M. P. J. (1994). Teaching evolution and natural selection: A look at textbooks and teachers. *Journal of Research in Science Teaching*, 31(5), 519–535.
- Allea. (2017). *The European Code of Conduct for Research Integrity*. <https://allea.org/code-of-conduct/>
- Amin, T. G., & Levrini, O. (2018). *Converging perspectives on conceptual change: Mapping an emerging paradigm in the learning sciences*. Routledge.
- Amin, T. G., Smith, C. L., & Wiser, M. (2014). Student conceptions and conceptual change: Three overlapping phases of research. In *Handbook of Research on Science Education, Volume II* (pp. 71–95). Routledge.
- Anderson, D. L., Fisher, K. M., & Norman, G. J. (2002). Development and evaluation of the conceptual inventory of natural selection. *Journal of Research in Science Teaching*, 39(10), 952–978.
- Anderson, T. R., Schönborn, K. J., du Plessis, L., Gupthar, A. S., & Hull, T. L. (2013). *Identifying and Developing Students' Ability to Reason with Concepts and Representations in Biology*.
- Appleton, L., & Mackie, J. (2019). Using Digital Organism Evolutionary Software in the Classroom. *The American Biology Teacher*, 81(1), 12–17.
- Ariew, A., Rohwer, Y., & Rice, C. (2017). Galton, reversion and the quincunx: The rise of statistical explanation. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 66, 63–72.
- Baalmann, W., Frerichs, V., Weitzel, H., Gropengießer, H., & Kattmann, U. (2004). Schülervorstellungen zu Prozessen der Anpassung – Ergebnisse einer Interviewstudie im Rahmen der Didaktischen Rekonstruktion. *Zeitschrift für Didaktik der Naturwissenschaften*, 10, 7–28.
- Bahar, M. (2003). Misconceptions in Biology Education and Conceptual Change Strategies. *Educational Sciences: Theory & Practice*, 3(1), 55.
- Baker, C. F., Fillmore, C. J., & Lowe, J. B. (1998). The berkeley framenet project. *36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics, Volume 1*, 86–90.
- Barnes, M. E., Evans, E. M., Hazel, A., Brownell, S. E., & Nesse, R. M. (2017). Teleological reasoning, not acceptance of evolution, impacts students' ability to learn natural selection. *Evolution: Education and*

- Outreach*, 10(1).
- Barradell, S. (2013). The identification of threshold concepts: A review of theoretical complexities and methodological challenges. *Higher Education*, 65(2), 265–276.
- Beggrow, E. P., & Nehm, R. H. (2012). Students' Mental Models of Evolutionary Causation: Natural Selection and Genetic Drift. *Evolution: Education and Outreach*, 5(3), 429–444.
- Bishop, B. A., & Anderson, C. W. (1990). Student conceptions of natural selection and its role in evolution. *Journal of Research in Science Teaching*, 27(5), 415–427.
- Borin, L., Dannélls, D., Forsberg, M., Gronostaj, M. T., & Kokkinakis, D. (2010). The past meets the present in Swedish FrameNet++. *14th EURALEX International Congress*, 269–281.
- Brenner, D. G., Matlen, B. J., Timms, M. J., Gochyyev, P., Grillo-Hill, A., Luttgen, K., & Varfolomeeva, M. (2017). Modeling Student Learning Behavior Patterns in an Online Science Inquiry Environment. *Technology, Knowledge and Learning*, 22(3), 405–425.
- Brumby, M. (1979). Problems in learning the concept of natural selection. *Journal of Biological Education*, 13(2), 119–122.
- Cardoso, J., Caetano, D., Abreu, R., Quadros, J., Santos, J. d., Ogasawara, E., & Lignani, L. (2020). Supporting the Learning of Evolution Theory Using an Educational Simulator. *IEEE Transactions on Learning Technologies*, 13(2), 417–424.
- Catley, K., Lehrer, R., & Reiser, B. (2005). Tracing a prospective learning progression for developing understanding of evolution. *Paper Commissioned by the National Academies Committee on Test Design for K-12 Science Achievement*, 67.
- Catley, K. M., & Novick, L. R. (2008). Seeing the Wood for the Trees: An Analysis of Evolutionary Diagrams in Biology Textbooks. *BioScience*, 58(10), 976–987.
- Ceballos, G., Ehrlich, P. R., Barnosky, A. D., García, A., Pringle, R. M., & Palmer, T. M. (2015). Accelerated modern human-induced species losses: Entering the sixth mass extinction. *Science Advances*, 1(5), e1400253.
- Chi, M. T. (2009). Three types of conceptual change: Belief revision, mental model transformation, and categorical shift. In *International handbook of research on conceptual change* (pp. 89–110). Routledge.
- Clarke-Midura, J., Pope, D. S., Maruca, S., Abraham, J. K., & Meir, E. (2018a). Iterative design of a simulation-based module for teaching evolution by natural selection. *Evolution: Education & Outreach*, 11(1), 1. edb.
- Clarke-Midura, J., Pope, D. S., Maruca, S., Abraham, J. K., & Meir, E. (2018b). Iterative design of a simulation-based module for teaching

- evolution by natural selection. *Evolution: Education and Outreach*, 11(1).
- Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design Experiments in Educational Research. *Educational Researcher*, 32(1), 9–13.
- Coulson, S. (2001). *Semantic leaps: Frame-shifting and conceptual blending in meaning construction*. Cambridge University Press.
- Cousin, G. (2006). An introduction to threshold concepts. *Planet*, 17, 4–5.
- Dasgupta, D., & Michalewicz, Z. (2013). *Evolutionary Algorithms in Engineering Applications*. Springer Science & Business Media.
- Demastes, S. S., Good, R. G., & Peebles, P. (1995). Students' conceptual ecologies and the process of conceptual change in evolution. *Science Education*, 79(6), 637–666.
- Demastes, S. S., Settlage, J., & Good, R. (1995). Students' conceptions of natural selection and its role in evolution: Cases of replication and comparison. *Journal of Research in Science Teaching*, 32(5), 535–550.
- Depew, D. J. (2016). Contingency, Chance, and Randomness in Ancient, Medieval, and Modern Biology. In G. Ramsey & C. H. Pence (Eds.), *Chance in evolution* (pp. 15–40). University of Chicago Press.
- Díaz, S., Settele, J., Brondízio, E., Ngo, H., Guèze, M., Agard, J., Arneth, A., Balvanera, P., Brauman, K., & Butchart, S. (2019). *Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services*.
- Dickes, A. C., & Sengupta, P. (2013). Learning Natural Selection in 4th Grade with Multi-Agent-Based Computational Models. *Research in Science Education*, 43(3), 921–953.
- Diffenbaugh, N. S., & Field, C. B. (2013). Changes in Ecologically Critical Terrestrial Climate Conditions. *Science*, 341(6145), 486–492.
- diSessa, A. A. (1993). Toward an Epistemology of Physics. *Cognition and Instruction*, 2/3, 105.
- Dobzhansky, T. (1973). Nothing in Biology Makes Sense except in the Light of Evolution. *The American Biology Teacher*, 35(3), 125–129.
- Driver, R., & Easley, J. (1978). Pupils and Paradigms: A Review of Literature Related to Concept Development in Adolescent Science Students. *Studies in Science Education*, 5(1), 61–84.
- Duit, R., Gropengießer, H., Kattmann, U., Komorek, M., & Parchmann, I. (2012). The Model of Educational Reconstruction – a Framework for Improving Teaching and Learning Science1. In D. Jorde & J. Dillon (Eds.), *Science Education Research and Practice in Europe* (pp. 13–37). Rotterdam. SensePublishers.
- Duit, R. H., & Treagust, D. F. (2012). Conceptual Change: Still a Powerful

- Framework for Improving the Practice of Science Instruction. In K. C. D. Tan & M. Kim (Eds.), *Issues and Challenges in Science Education Research: Moving Forward* (pp. 43–54). Springer Netherlands.
- Duncan, R. G., & Reiser, B. J. (2007). Reasoning across ontologically distinct levels: Students' understandings of molecular genetics. *Journal of Research in Science Teaching*, 44(7), 938–959.
- Duret, L. (2008). *Neutral Theory: The Null Hypothesis of Molecular Evolution*.
- Erickson, F. (2012). Qualitative Research Methods for Science Education. In B. J. Fraser, K. Tobin, & C. J. McRobbie (Eds.), *Second International Handbook of Science Education* (pp. 1451–1469). Springer Netherlands.
- Evans. (2000). The Emergence of Beliefs About the Origins of Species in School-Age Children. *Merrill-Palmer Quarterly*, 46(2), 221–254. JSTOR.
- Evans. (2001). Cognitive and contextual factors in the emergence of diverse belief systems: Creation versus evolution. *Cognitive Psychology*, 42(3), 217–266.
- Evans, E. M. (2008). Conceptual change and evolutionary biology: A developmental analysis. *International Handbook of Research on Conceptual Change*, 263–294.
- Evans, V. (2019). *Cognitive Linguistics: A Complete Guide*. Edinburgh University Press.
- Everitt, B., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster Analysis*. Hoboken : Wiley, 2011.
- Fauconnier, G., & Turner, M. (1998). Conceptual Integration Networks. *Cognitive Science*, 22(2), 133–187.
- Federer, M. R., Nehm, R. H., Opfer, J. E., & Pearl, D. (2015). Using a Constructed-Response Instrument to Explore the Effects of Item Position and Item Features on the Assessment of Students' Written Scientific Explanations. *Research in Science Education*, 45(4), 527–553.
- Ferrari, M., & Chi, M. T. (1998). The nature of naive explanations of natural selection. *International Journal of Science Education*, 20(10), 1231–1256.
- Fillmore, C. J. (1982). Frame semantics. In *Linguistics in the Morning Calm*. Hanshin Publishing Co.
- Flanagan, M. (2015, April 16). *The Threshold Concept*. <http://www.ee.ucl.ac.uk/~mflanaga/thresholds.html>
- Forbes, M. L. (1978). Simulation of Natural Selection on the Programmable Calculator. *Journal of College Science Teaching*, 8(2), 95. edsjsr.
- Gale, N. K., Heath, G., Cameron, E., Rashid, S., & Redwood, S. (2013). Using the framework method for the analysis of qualitative data in

- multi-disciplinary health research. *BMC Medical Research Methodology*, 13(1), 117.
- Garvin-Doxas, K., & Klymkowsky, M. W. (2008). Understanding randomness and its impact on student learning: Lessons learned from building the Biology Concept Inventory (BCI). *CBE-Life Sciences Education*, 7(2), 227–233.
- Gilbert, J. K. (2008). Visualization: An emergent field of practice and enquiry in science education. In *Visualization: Theory and practice in science education* (pp. 3–24). Springer.
- Göransson, A., Orraryd, D., Fiedler, D., & Tibell, L. A. E. (2020). Conceptual Characterization of Threshold Concepts in Student Explanations of Evolution by Natural Selection and Effects of Item Context. *CBE—Life Sciences Education*, 19(1), ar1.
- Gregory, T. R. (2009). Understanding Natural Selection: Essential Concepts and Common Misconceptions. *Evolution: Education and Outreach*, 2(2), 156–175.
- Ha, M., & Nehm, R. H. (2013). Darwin's Difficulties and Students' Struggles with Trait Loss: Cognitive-Historical Parallelisms in Evolutionary Explanation. *Science & Education*, 23(5), 1051–1074.
- Hamid, J. S., Meaney, C., Crowcroft, N. S., Granerod, J., & Beyene, J. (2010). Cluster analysis for identifying sub-groups and selecting potential discriminatory variables in human encephalitis. *BMC Infectious Diseases*, 10(1), 364.
- Harlen, W. (2010). *Principles and big ideas of science education*. Association for science education.
- Harmon, K. (2011). *Evolution Abroad: Creationism Evolves in Science Classrooms around the Globe*. Scientific American. <https://www.scientificamerican.com/article/evolution-education-abroad/>
- Hayes, A. F., & Krippendorff, K. (2007). Answering the Call for a Standard Reliability Measure for Coding Data. *Communication Methods and Measures*, 1(1), 77–89.
- Heddy, B. C., & Sinatra, G. M. (2013). Transforming Misconceptions: Using Transformative Experience to Promote Positive Affect and Conceptual Change in Students Learning About Biological Evolution: TRANSFORMING MISCONCEPTIONS. *Science Education*, 97(5), 723–744.
- Heitz, J. G., Cheetham, J. A., Capes, E. M., & Jeanne, R. L. (2010). Interactive Evolution Modules Promote Conceptual Change. *Evolution: Education and Outreach*, 3(3), 436–442.
- Hendry, A. P., Kinnison, M. T., Heino, M., Day, T., Smith, T. B., Fitt, G., Bergstrom, C. T., Oakeshott, J., Jørgensen, P. S., Zalucki, M. P., Gilchrist, G., Southerton, S., Sih, A., Strauss, S., Denison, R. F., & Carroll, S. P. (2011). Evolutionary principles and their practical

- application. *Evolutionary Applications*, 4(2), 159–183.
- Heredia, S. C., Furtak, E. M., & Morrison, D. (2016). Exploring the influence of plant and animal item contexts on student response patterns to natural selection multiple choice items. *Evolution: Education and Outreach*, 9(1), 10.
- Herrero, D., Del Castillo, H., Monjelat, N., García-Varela, A. B., Checa, M., & Gomez, P. (2014). Evolution and natural selection: Learning by playing and reflecting. *Journal of New Approaches in Educational Research*, 3(1), 26–33.
- Hodge, J. (2016). Chance and Chances in Darwin's Early Theorizing and in Darwinian Theory Today. In G. Ramsey & C. H. Pence (Eds.), *Chance in evolution* (pp. 41–75). University of Chicago Press.
- Hodgson, J. Y. S. (2019). A Simplified Bioenergetics Model of a Hammerhead Shark for Teaching Natural Selection. *American Biology Teacher (University of California Press)*, 81(2), 115. edb.
- Hoekstra, H. E. (2011). *From Darwin to DNA: The Genetic Basis of Color Adaptations*. Roberts and Company Publishers. <https://dash.harvard.edu/handle/1/34728587>
- Holley, J. W., & Guilford, J. P. (1964). A note on the G index of agreement. *Educational and Psychological Measurement*, 24(4), 749–753.
- Hrepic, Z., Zollman, D. A., & Rebello, N. S. (2010). Identifying students' mental models of sound propagation: The role of conceptual blending in understanding conceptual change. *Physical Review Special Topics - Physics Education Research*, 6(2), 020114.
- Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277–1288.
- Hunter, J. (1958). Teaching Evolution to High School Students. *The American Biology Teacher*, 20(8), 283–284.
- Islam, M. M., Valderas, J. M., Yen, L., Dawda, P., Jowsey, T., & McRae, I. S. (2014). Multimorbidity and Comorbidity of Chronic Diseases among the Senior Australians: Prevalence and Patterns. *PLoS ONE*, 9(1), e83783.
- Jakobi, S. R. (2010). "Little Monkeys on the Grass..." How People for and Against Evolution Fail to Understand the Theory of Evolution. *Evolution: Education and Outreach*, 3(3), 416–419.
- Jensen, M. S., & Finley, F. N. (1995). Teaching evolution using historical arguments in a conceptual change strategy. *Science Education*, 79(2), 147–166.
- Johnson, C. N. (2015). *Darwin's dice: The idea of chance in the thought of Charles Darwin*. Oxford University Press.
- Johnstone, A. H. (1991). Why is science difficult to learn? Things are seldom what they seem. *Journal of Computer Assisted Learning*, 7(2), 75–83.

- Jones, T. C., & Laughlin, T. F. (2010). PopGen Fishbowl: A Free Online Simulation Model of Microevolutionary Processes. *American Biology Teacher (National Association of Biology Teachers)*, 72(2), 100–103.
- Jördens, J., Asshoff, R., Kullmann, H., & Hammann, M. (2016). Providing vertical coherence in explanations and promoting reasoning across levels of biological organization when teaching evolution. *International Journal of Science Education*, 38(6), 960–992.
- Kalinowski, S. T., Leonard, M. J., & Andrews, T. M. (2010). Nothing in evolution makes sense except in the light of DNA. *CBE-Life Sciences Education*, 9(2), 87–97.
- Kampourakis, K., & Zogza, V. (2008). Preliminary Evolutionary Explanations: A Basic Framework for Conceptual Change and Explanatory Coherence in Evolution. *Science & Education*, 18(10), 1313–1340.
- Kattmann, U. (2008). Learning biology by means of anthropomorphic conceptions. *Biology in Context: Learning and Teaching for the 21st Century*. London, UK: Institute of Education, University of London.
- Kattmann, U., Duit, R., Gropengiesser, H., & Komorek, M. (1996). *Educational Reconstruction – Bringing together Issues of scientific clarification and students' conceptions*. NARST 1996.
- Keeler, C. E. (1941). Progressive Diagrams for Teaching Biology. *The American Biology Teacher*, 3(7), 227–232.
- Kersting, M., Henriksen, E. K., Bøe, M. V., & Angell, C. (2018). General relativity in upper secondary school: Design and evaluation of an online learning environment using the model of educational reconstruction. *Physical Review Physics Education Research*, 14(1), 010130.
- Kersting, M., & Steier, R. (2018). Understanding Curved Spacetime. *Science & Education*, 27(7), 593–623.
- Kimura. (1991). The neutral theory of molecular evolution: A review of recent evidence. *Idengaku Zasshi*, 66(4), 367–386.
- Kimura, Motoo. (1968). Evolutionary Rate at the Molecular Level. *Nature*, 217(5129), 624–626
- Klafki, W. (1995). Didactic analysis as the core of preparation of instruction (Didaktische Analyse als Kern der Unterrichtsvorbereitung). *Journal of Curriculum Studies*, 27(1), 13–30.
- Krippendorff, K. (2013). *Content analysis: An introduction to its methodology*. Thousand Oaks, Calif. ; London : SAGE, 2013.
- Kuschmierz, P., Meneganzin, A., Pinxten, R., Pievani, T., Cvetković, D., Mavrikaki, E., Graf, D., & Beniermann, A. (2020). Towards common ground in measuring acceptance of evolution and knowledge about evolution across Europe: A systematic review of the state of research. *Evolution: Education and Outreach*, 13(1), 18.
- Lewontin, R. C. (1970). The Units of Selection. *Annual Review of*

- Ecology & Systematics*, 1, 1–18.
- Lin, T.-J., Lin, T.-C., Potvin, P., & Tsai, C.-C. (2019). Research trends in science education from 2013 to 2017: A systematic content analysis of publications in selected journals. *International Journal of Science Education*, 41(3), 367–387.
- Lind Pantzare, A., & Abrahamsson, M. (2015). *Ämnesproven i grundskolans årskurs 9 och specialskolans årskurs 10*. 27.
- Liu, S., Shen, J., Fang, S., Li, K., Liu, J., Yang, L., Hu, C.-D., & Wan, J. (2020). Genetic Spectrum and Distinct Evolution Patterns of SARS-CoV-2. *Frontiers in Microbiology*, 11.
- Luria, S. E., & Delbrück, M. (1943). Mutations of Bacteria from Virus Sensitivity to Virus Resistance. *Genetics*, 28(6), 491–511.
- Malone, K. L., Schunn, C. D., & Schuchardt, A. M. (2018). Improving Conceptual Understanding and Representation Skills Through Excel-Based Modeling. *JOURNAL OF SCIENCE EDUCATION AND TECHNOLOGY*, 27(1), 30–44.
- Marbach-Ad, G., & Stavy, R. (2000). Students' cellular and molecular explanations of genetic phenomena. *Journal of Biological Education*, 34(4), 200–205.
- Maxwell, J. A. (2012). *Qualitative research design: An interactive approach*. SAGE Publications.
- Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology*, 90(2), 312.
- Mayr, E. (1982). *The growth of biological thought: Diversity, evolution, and inheritance*. Belknap P. of Harvard U.P.
- Mayr, E. (1991). *One long argument: Charles Darwin and the genesis of modern evolutionary thought*. Harvard Univ. Press.
- Mead, L. S., & Scott, E. C. (2010). Problem Concepts in Evolution Part II: Cause and Chance. *Evolution: Education and Outreach*, 3(2), 261–264.
- Meester, L. D., Stoks, R., & Brans, K. I. (2018). Genetic adaptation as a biological buffer against climate change: Potential and limitations. *Integrative Zoology*, 13(4), 372–391. <https://doi.org/10.1111/1749-4877.12298>
- Meyer, J., & Land, R. (2003). Threshold concepts and troublesome knowledge: Linkages to ways of thinking and practising within the disciplines. In C. Rust (Ed.), *Improving Student Learning—Theory and Practice Ten Years On*. (pp. 412–424). Oxford: Oxford Centre for Staff and Learning Development (OCSLD).
- Milo, R., & Phillips, R. (2015). *Cell Biology by the Numbers*. <http://book.bionumbers.org/>
- Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia

- learning: The role of modality and contiguity. *Journal of Educational Psychology*, 91(2), 358.
- Nehm, R. H., Ha, M., Rector, M., Opfer, J., Perrin, L., Ridgway, J., & Mollohan, K. (2010). *Scoring guide for the open response instrument (ORI) and evolutionary gain and loss test (EGALT)*. Technical Report of National Science Foundation REESE Project 0909999. Accessed online 10 Jan 2011 at: <http://evolutionassessment.org>. http://www.academia.edu/download/30469189/technical_report_scoring_guide_2010_1.pdf
- Nehm, & Reilly, L. (2007). Biology majors' knowledge and misconceptions of natural selection. *BioScience*, 57(3), 263–272.
- Nehm, Ross H., & Ha, M. (2011). Item feature effects in evolution assessment. *Journal of Research in Science Teaching*, 48(3), 237–256.
- Nehm, Ross H., Poole, T. M., Lyford, M. E., Hoskins, S. G., Carruth, L., Ewers, B. E., & Colberg, P. J. (2009). Does the segregation of evolution in biology textbooks and introductory courses reinforce students' faulty mental models of biology and evolution? *Evolution: Education and Outreach*, 2(3), 527–532.
- Nehm, Ross H., Rector, M. A., & Ha, M. (2010). “Force-Talk” in Evolutionary Explanation: Metaphors and Misconceptions. *Evolution: Education and Outreach*, 3(4), 605–613.
- Nehm, R. H. & Schonfeld, I. S. (2008). Measuring knowledge of natural selection: A comparison of the CINS, an open-response instrument, and an oral interview. *Journal of Research in Science Teaching*, 45(10), 1131–1160.
- NGSS Lead States. (2013). *Next Generation Science Standards: For States, By States*. <https://www.nextgenscience.org/>
- Niebert, K., Marsch, S., & Treagust, D. F. (2012). Understanding needs embodiment: A theory-guided reanalysis of the role of metaphors and analogies in understanding science. *Science Education*, 96(5), 849–877.
- Ortiz-Cordova, A., Yang, Y., & Jansen, B. J. (2015). External to internal search: Associating searching on search engines with searching on sites. *Information Processing & Management*, 51(5), 718–736.
- Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive Load Theory and Instructional Design: Recent Developments. *Educational Psychologist*, 38(1), 1–4.
- Paz-y-Miño C., G., Espinosa, A., & Bai, C. Y. (2011). The Jackprot Simulation Couples Mutation Rate with Natural Selection to Illustrate How Protein Evolution Is Not Random. *EVOLUTION EDUCATION AND OUTREACH*, 3, 502. edsbl.
- Pigliucci, M., & Kaplan, J. M. (2006). *Making sense of evolution. [Elektronisk resurs]: The conceptual foundations of evolutionary*

- biology* (INTERNET INTERNET). Chicago : University of Chicago Press, 2006.
- Pigliucci, M., Müller, G., & Konrad Lorenz Institute for Evolution and Cognition Research (Eds.). (2010). *Evolution, the extended synthesis*. MIT Press.
- Pope, D. S., Rounds, C. M., & Clarke-Midura, J. (2017). Testing the effectiveness of two natural selection simulations in the context of a large-enrollment undergraduate laboratory class. *Evolution: Education and Outreach*, 10(1).
- Portin, P., & Wilkins, A. (2017). The Evolving Definition of the Term “Gene.” *Genetics*, 205(4), 1353–1364. <https://doi.org/10.1534/genetics.116.196956>
- Powell, T. E., Boomgaarden, H. G., Swert, K. D., & Vreese, C. H. de. (2015). A Clearer Picture: The Contribution of Visuals and Text to Framing Effects. *Journal of Communication*, 65(6), 997–1017. <https://doi.org/10.1111/jcom.12184>
- Protas, M., & Jeffery, W. R. (2012). Evolution and development in cave animals: From fish to crustaceans: Evolution and development in cave animals. *Wiley Interdisciplinary Reviews: Developmental Biology*, 1(6), 823–845.
- Pugh, K. J., Koskey, K. L. K., & Linnenbrink-Garcia, L. (2014). High school biology students’ transfer of the concept of natural selection: A mixed-methods approach. *Journal of Biological Education*, 48(1), 23–33.
- Reeves, T. D., & Marbach-Ad, G. (2016). Contemporary Test Validity in Theory and Practice: A Primer for Discipline-Based Education Researchers. *CBE—Life Sciences Education*, 15(1), rm1.
- Robson, C. (2011). *Real world research: A resource for users of social research methods in applied settings*. Chichester : Wiley, 2011.
- Rominger, C., Weiss, E. M., Fink, A., Schuler, G., & Papousek, I. (2011). Allusive thinking (cognitive looseness) and the propensity to perceive “meaningful” coincidences. *Personality and Individual Differences*, 51(8), 1002–1006.
- Ross, P. M., Taylor, C. E., Hughes, C., Whitaker, N., Lutze-Mann, L., Kofod, M., & Tzioumis, V. (2010). Threshold concepts in learning biology and evolution. *Biology International*, 47, 47–54.
- Salwén, H. (2019). Threshold concepts, obstacles or scientific dead ends? *Teaching in Higher Education*, 1–14.
- Schramm, T., Schachtschneider, Y., & Schmiemann, P. (2019). Understanding the tree of life: An overview of tree-reading skill frameworks. *Evolution: Education and Outreach*, 12(1).
- Shtulman, A. (2006). Qualitative differences between naïve and scientific theories of evolution. *Cognitive Psychology*, 52(2), 170–194.

- Siegel, S., & Castellan, N. J., Jr. (1988). *Nonparametric statistics for the behavioral sciences* (Campus Norrköpings bibliotek - Boksamlingen 519). New York : McGraw-Hill, 1988.
- Sinatra, G. M., Brem, S. K., & Evans, E. M. (2008). Changing Minds? Implications of Conceptual Change for Teaching and Learning about Biological Evolution. *Evolution: Education and Outreach*, 1(2), 189–195.
- Skolverket. (2013). *Curriculum for the upper secondary school*. <https://www.skolverket.se/getFile?file=2975>
- Smetana, L. K., & Bell, R. L. (2012). Computer Simulations to Support Science Instruction and Learning: A critical review of the literature. *International Journal of Science Education*, 34(9), 1337–1370.
- Smith, M. U. (2009a). Current Status of Research in Teaching and Learning Evolution: I. Philosophical/Epistemological Issues. *Science & Education*, 19(6–8), 523–538.
- Smith, M. U. (2009b). Current Status of Research in Teaching and Learning Evolution: II. Pedagogical Issues. *Science & Education*, 19(6–8), 539–571.
- Stearns, S. C. (2012). Evolutionary medicine: Its scope, interest and potential. *Proceedings of the Royal Society B: Biological Sciences*, 279(1746), 4305–4321.
- Stebbins, R. C., & Allen, B. (1975). Simulating Evolution. *The American Biology Teacher*, 37(4), 206–211. JSTOR. <https://doi.org/10.2307/4445177>
- Stigler, S. M. (2010). Darwin, Galton and the Statistical Enlightenment. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 173(3), 469–482.
- Strike, K. A., & Posner, G. J. (1982). Conceptual change and science teaching. *European Journal of Science Education*, 4(3), 231–240.
- Taber, K. S. (2014). Alternative Conceptions/Frameworks/Misconceptions. In R. Gunstone (Ed.), *Encyclopedia of Science Education* (pp. 1–5). Springer Netherlands.
- Talanquer, V. (2007). Explanations and Teleology in Chemistry Education. *International Journal of Science Education*, 29(7), 853–870.
- Tibell, L. A. E., & Harms, U. (2017). Biological Principles and Threshold Concepts for Understanding Natural Selection. *Science & Education*, 1–21.
- Tight, M. (2014). Theory development and application in higher education research: The case of threshold concepts. In J. Huisman & M. Tight (Eds.), *International Perspectives on Higher Education Research* (Vol. 10, pp. 249–267). Emerald Group Publishing Limited.
- Treagust, D., Won, M., & Duit, R. (2014). Paradigms in Science Education Research. In *Handbook of Research on Science Education* (pp.

- 3–17). Routledge.
- Tshuma, T., & Sanders, M. (2015). Textbooks as a Possible Influence on Unscientific Ideas about Evolution. *Journal of Biological Education*, 49(4), 354–369.
- Tsui, C.-Y., & Treagust, D. F. (2013). Introduction to multiple representations: Their importance in biology and biological education. In *Multiple representations in biological education* (pp. 3–18). Springer.
- Turner, M., & Fauconnier, G. (1995). Conceptual Integration and Formal Expression. *Metaphor & Symbolic Activity*, 10(3), 183.
- Tversky, B. (2011). Visualizing Thought. *Topics in Cognitive Science*, 3(3), 499–535.
- Tversky, B., Morrison, J. B., & Betrancourt, M. (2002). Animation: Can it facilitate? *International Journal of Human-Computer Studies*, 57(4), 247–262.
- United Nations. (2015). *Transforming Our World: The 2030 Agenda for Sustainable Development*. UN Publishing, New York.
- Vollmer, G. (1984). Mesocosm and Objective Knowledge. In F. M. Wuketits (Ed.), *Concepts and Approaches in Evolutionary Epistemology* (pp. 69–121). Springer Netherlands.
- Vosniadou, S. (2013a). Conceptual Change In Learning and Instruction: The Framework Theory Approach. In *International handbook of research on conceptual change* (pp. 23–42). Routledge.
- Vosniadou, S. (2013b). *International Handbook of Research on Conceptual Change*. Routledge.
- Vosniadou, S. (2019). The Development of Students' Understanding of Science. *Frontiers in Education*, 4, 32.
- Vosniadou, S., & Ioannides, C. (1998). From conceptual development to science education: A psychological point of view. *International Journal of Science Education*, 20(10), 1213–1230.
- Wanford, J., Aidley, J., Bayliss, C., Ketley, J., & Goodwin, M. (2018). Simulating phase variation: A practical approach to teaching mutation and diversity*. *Journal of Biological Education (Routledge)*, 52(1), 47–53. a9h.
- Weiskopf, S. R., Rubenstein, M. A., Crozier, L. G., Gaichas, S., Griffis, R., Halofsky, J. E., Hyde, K. J. W., Morelli, T. L., Morissette, J. T., Muñoz, R. C., Pershing, A. J., Peterson, D. L., Poudel, R., Staudinger, M. D., Sutton-Grier, A. E., Thompson, L., Vose, J., Weltzin, J. F., & Whyte, K. P. (2020). Climate change effects on biodiversity, ecosystems, ecosystem services, and natural resource management in the United States. *Science of The Total Environment*, 733, 137782.
- Winther, R. G. (2000). Darwin on Variation and Heredity. *Journal of the History of Biology*, 33(3), 425–455.
- Xiang, L. (2011). *A Collective Case Study of Secondary Students' Model-*

- Based Inquiry on Natural Selection through Programming in an Agent-Based Modeling Environment*. ProQuest LLC. Ann Arbor, MI.
- Xu, S., & Lorber, M. F. (2014). Interrater agreement statistics with skewed data: Evaluation of alternatives to Cohen's kappa. *Journal of Consulting and Clinical Psychology*, 82(6), 1219–1227.
- Yamanoi, T., & Iwasaki, W. M. (2015). Origami Bird Simulator: A Teaching Resource Linking Natural Selection and Speciation. *Evolution: Education and Outreach*, 8(1).
- Ziadie, M. A., & Andrews, T. C. (2018). Moving Evolution Education Forward: A Systematic Analysis of Literature to Identify Gaps in Collective Knowledge for Teaching. *CBE—Life Sciences Education*, 17(1), ar11.

Appendices

I. Paper I unpublished data

127 visualizations were analyzed by two raters for each visualization. 79 of the 127 (62%) that contained at least on threshold concept in the first analysis were analyzed in detail regarding the mode of concept expression (visual or oral). The results are presented in the table below. (The analysis in Paper I was limited to a subset of these 127 visualizations that were videos, excluding simulations, games etc.).

<i>Threshold concept</i>	<i>Oral</i>	<i>Visual</i>	(% of visualizations)			Total
			<i>Oral and vis.</i>	<i>Only oral</i>	<i>Only visual</i>	
Temporal scale	31	35	19	13	15	46
Temporal scale indica- tor	1	19	1	0	19	20
Spatial scale	16	20	12	2	8	22
Spatial scale indicator	2	3	0	2	3	6
Probability	6	6	4	2	2	9
Randomness	13	3	1	13	2	16
Total	46	54	36	9	18	64

II. Consent form and information to participants

Example of consent form and information to participants.

Information till elever och vårdnadshavare

Digitala simuleringar av biologi

Hej!

Mitt namn är Andreas Göransson och jag är doktorand vid Linköpings Universitet. Jag och mina forskarkollegor driver ett forskningsprojekt för att undersöka hur man kan underlätta förståelsen av biologi med hjälp av datorsimuleringar.

Denna studie syftar till att undersöka hur elever förstår biologi med hjälp av datorsimuleringar. Undersökningen är viktig för att kunna utveckla bättre undervisning och läromedel. Du kommer att få fylla i en kort enkät med frågor om biologi före och efter undersökningen. Du kommer även att få jobba med några datorsimuleringar och diskutera ett antal frågor tillsammans med klasskamraterna. Vi vill även spela in era diskussioner kring simuleringen på video för att senare kunna titta på hur ni förstod simuleringarna.

Deltagandet är helt frivilligt och du kan när som helst under tiden undersökningen pågår meddela att du inte vill fortsätta. Genom att delta i undersökningen lämnar du ditt samtycke till att enkäterna används för forskning enligt beskrivningen nedan. Alla data samlas in helt anonymt. De insamlade uppgifterna kommer endast att användas i forskningssyfte och främst för utveckling av en visualisering.

De insamlade uppgifterna kan komma att ingå i bearbetad form i vetenskapliga publikationer. Om du är intresserad av att ta del av resultatet vid eventuell publicering kan du vända till dig till ansvarig för undersökningen.

Syfte och genomförande

Huvudsyftet är att undersöka på vilka sätt olika tekniker och exempel kan stödja elever i sin förståelse av biologi. Vi vill också undersöka hur digital teknik kan gestalta abstrakta begrepp och hur de på detta sätt kan göras tydliga och förståeliga för elever. Studien kommer genomföras som en del av undervisningen och äga rum i skolans befintliga klassrum. Elever som deltar kommer få svara på skriftliga frågor, arbeta med en gruppuppgift och arbeta med datorsimuleringar. Ett fåtal enskilda intervjuer kan komma att genomföras vid ett senare tillfälle. Totalt kommer undersökningen inte ta mer än två timmar.

Frivilligt deltagande

För att kunna medverka i studien behöver du (och även din målsman om du är under 16 år) fylla i ett deltagandeavtal på baksidan av detta brev. Deltagande i den här studien är helt frivilligt och man kan när som helst välja att avbryta sitt deltagande.

Tack för din medverkan!

Andreas Göransson, doktorand vid Institutionen för Teknik och Naturvetenskap
Linköpings universitet, 601 74, Norrköping

andreas.c.goransson@liu.se

Deltagandeavtal

Detta avtal är ett medgivande till materialanvändning från den som deltagit i en enkätundersökning, en intervju eller vid en bild och/eller ljudupptagning med avsikt att ge råmaterial till forskning vid institutionen ITN på Linköpings universitet.

Användande av materialet

Användande av materialet avser t.ex. analys av deltagarnas interaktion med såväl varandra som maskin- och mjukvara. Råmaterialet kommer att förvaras på ett säkert sätt och ej spridas utanför den inblandade forskargruppen.

- ☐ Jag medger att delta i studien vilket innebär att enkät- och intervjumaterial, upptaget bild- och ljudmaterial där jag medverkar samt det material jag varit med och skapat under projektet får användas i *forskningssyfte*, förutsatt att det hanteras i enighet med vedertagen svensk forskningsetik.
- ☐ Jag medger ej

Presentation av resultat från studie

- Text Ja ☐ Nej ☐ Jag medger att enkät- och intervjumaterial såsom aidentifierade och anonymiserade citat från video och ljudupptagningar får användas i samband med *presentationer* och *publicering* av studiens resultat.
- Bilder Ja ☐ Nej ☐
- Video Ja ☐ Nej ☐
- Ljud Ja ☐ Nej ☐

Ångerrätt

Jag kan när jag vill ändra mina nuvarande medgivanden och avbryta mitt deltagande i studien. Förändringar i avtalet blir giltiga när de tas emot av kontaktpersonen och gäller inte retroaktivt avseende publicering utförd enligt tidigare avtal.

Underskrift

Namn (textat)	Personnummer
Ort och datum	
Underskrift	
Målsmans underskrift (för den som är yngre än 16 år)	

Kontaktperson: Namn: Andreas Göransson / Lena Tibell
E-post: andreas.c.goransson@liu.se / lena.tibell@liu.se
Adress: ITN, Campus Norrköping, SE-601 74 Norrköping

Papers

The papers associated with this thesis have been removed for copyright reasons. For more details about these see:

<http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-173337>

FACULTY OF SCIENCE AND ENGINEERING

Linköping Studies in Science and Technology, Dissertation No. 2122, 2021
Department of Science and Technology

Linköping University
SE-581 83 Linköping, Sweden

www.liu.se