KØBENHAVNS UNIVERSITET
INSTITUT FOR NATURFAGENES DIDAKTIK

Independence in Secondary Probability and Statistics: Content Analysis and Task Design

Jonas Uglebjerg

Master's Thesis –Math

Advisor: Carl Winsløw
December 20th, 2022

IND's studenterserie nr. 112, 2023

INSTITUT FOR NATURFAGENES DIDAKTIK, www.ind.ku.dk
Alle publikationer fra IND er tilgængelige via hjemmesiden.

IND's studenterserie

70. Jakob Holm: The Implementation of Inquiry-based Teaching (2019)
71. Louise Uglebjerg: A Study and Research Path (2019)
72. Anders Tørring Kolding & Jonas Tarp Jørgensen: Physical Activity in the PULSE Exhibit (2019)
73. Simon Arent Vedel: Teaching the Formula of Centripetal Acceleration (2019)
74. Aputsiaq Bent Simonsen: Basic Science Course (NV) (2019)
75. Svenning Helth Møller: Peer-feedback (2019)
76. Lars Hansen & Lisbeth Birch Jensen: Feedbackformater på Mulernes Legatskole (2019)
77. Kirsi Inkeri Pakkanen: Autobiographical narratives with focus on science (2019)
78. Niels Jacob Jensen: Engineering i naturen og på naturskolen (2019)
79. Yvonne Herguth Nygaard: Diskursanalyse af litteraturen og hos lærer i forbindelse med brugen af eksterne læringsmiljø, med en underviser tilknyttet (2019)
80. Trine Jørgensen: Medborgerskab i naturfagsundervisningen på KBHSYD (2019)
81. Morten Terp Randrup: Dannelse i Fysik C (2019)
82. Thomas Mellergaard Amby: Undersøgelses baseret naturfagsundervisning og science writing heuristic (2019)
83. Freja Elbro: Important prerequisites to understanding the definition of limit (2019)
84. Mathilde Sexauer Bloch Kloster: Inquiry-Based Science Education (IBSE) (2019)
85. Casper Borup Frandsen: Undersøgelsesbaseret undervisning i idrætsundervisningen på gymnasieskolen (2019)
86. Vibeke Ankjer Vestermarken: An Inquiry Based Introduction to Binomial Distributions (2019)
87. Jesper Jul Jensen: Formativ evaluering og faglige samspil i almen studieforberedelse (2020)
88. Karen A. Voigt: Assessing Student Conceptions with Network Theory - Investigating Student Conceptions in the Force Concept Inventory Using MAMCR (2020)
89. Julie Hougaard Overgaard: Using virtual experiments as a preparation for large scale facility experiments (2020)
90. Maria Anagnostou: Trigonometry in upper secondary school context: identities and functions (2020)
91. Henry James Evans: How Do Different Framings Of Climate Change Affect Pro-environmental Behaviour? (2020)
92. Mette Jensen: Study and Research Paths in Discrete Mathematics (2020)
93. Jesper Hansen: Effekten og brugen af narrative læringsspil og simuleringer i gymnasiet (2020)
94. Mie Haumann Petersen: Bilingual student performance in the context of probability and statistics teaching in Danish High schools (2020)
95. Caroline Woergaard Gram: "Super Yeast" - The motivational potential of an inquiry-based experimental exercise (2021)
96. Lone Brun Jakobsen: Kan man hjælpe elevers forståelse af naturvidenskab ved at lade dem formulere sig om et naturvidenskabeligt emne i et andet fag? (2021)
97. Maibritt Oksen og Morten Kjøller Hegelund: Styrkelse af motivation gennem Webinar og Green Screen (2021)
98. Søren Bystrup Jacobsen: Peer feedback: Fra modstand til mestring? (2021)
99. Bente Guldbrandsen: Er der nogen, som har spurgt en fysiklærer? (2021)
100. Iben Vernegren Christensen: Bingoplader i kemiundervisningen – en metode til styrkelse af den faglige samtale? (2021)
101. Claus Axel Frimann Kristinson Bang: Probability, Combinatorics, and Lesson Study in Danish High School (2021)
102. Derya Diana Cosan: A Diagnostic Test for Danish Middle School Arithmetics (2021)
103. Kasper Rytter Falster Dethlefsen: Formativt potentiale og udbytte i Structured Assessment Dialogue (2021)
104. Nicole Jonasen: A diagnostic study on functions (2021)
105. Trine Nørgaard Christensen: Organisatorisk læring på teknisk eux (2021)
106. Simon Funch: Åben Skole som indgang til tværfagligt samarbejde (2022)
107. Hans-Christian Borggreen Keller: Stem som interdisciplinær undervisningsform (2022)
108. Marie-Louise Krarup, Jakob Holm Jakobsen, Michelle Kyk & Malene Hermann Jensen: Implementering af STEM i grundskolen (2022)
109. Anja Rousing Lauridsen & Jonas Traczyk Jensen: Grundskoleelevers oplevelse af SSI-undervisning i en STEM-kontekst. (2022)
110. Aurora Olden Aglen: Danish upper secondary students' apprehensions of the equal sign (2023)
111. Metine Rahbek Tarp & Nicolaj Pape Frantzen: Machine Learning i gymnasiet (2023)
112. **Jonas Uglebjerg: Independence in Secondary Probability and Statistics: Content Analysis and Task Design (2023)**

***IND's studenterserie** omfatter kandidatspecialer, bachelorprojekter og masterafhandlinger skrevet ved eller i tilknytning til Institut for Naturfagenes Didaktik. Disse drejer sig ofte om uddannelsesfaglige problemstillinger, der har interesse også uden for universitetets mure. De publiceres derfor i elektronisk form, naturligvis under forudsætning af samtykke fra forfatterne. Det er tale om studenterarbejder, og ikke endelige forskningspublikationer.*
Se hele serien på: www.ind.ku.dk/publikationer/studenterserien/

Abstract

This thesis is written on the subject of Mathematical Didactics. The intention of this thesis is to examine how conditional probability and independence should be taught such that students acquire solid knowledge about the subject that is conceptualized in a non-misleading way. This thesis contains a presentation of the place of conditional probability and independence in curricula and textbooks today and within the past 15 years. This is followed by a review of preliminary theory on how misconception can be avoided in conditional probability and independence. This review is used to design a teaching course of five lessons of 55 minutes, that aims to prevent students from forming misconceptions and to increase students' linguistic and formalistic abilities. This didactic design is made within the framework of the Theory of Didactic Situations, which is also explained. The developed design has been tested on a second-year class in a Danish upper secondary school. The thesis contains an a priori analysis, which was made before this design was tested, and an a posteriori analysis of the collected data and observations. Based on this, it can be concluded that it takes more than five lessons to deconceptualize the students' misconceptions. It can also be concluded that it is possible to increase the students' linguistic abilities, but it requires changes in the teaching design and a new study to be able to conclude that the students' formalistic abilities can be improved.

Acknowledgements

First of all, I would like to thank Carl Winsløw for advising me with great calmness, care and accessibility. Receiving calm and precise feedback is of great value when you are a stressed thesis writer.

I would also like to say a big thank you to my Ida-Marie, for being there for me every single day. Whenever I needed comfort, she was ready with hugs and words of encouragement. I am looking forward to spending my weekends with you, instead of my thesis.

My father has been an important and competent sparring partner in relation to teaching, as he was able to contribute with insightful considerations based on his many years of teaching experience. In addition, I want to thank you for the extensive proofreading.

I want to thank my mother as well. Thank you for being supportive when I did not believe I was able to finish my thesis and for encouraging me to keep going. In addition, I especially want to thank you for driving all the way to Copenhagen to pick me up when I needed to come home and my the support from my family.

Last but not least, a thank you to my two sisters, Louise and Marie. You have both (in very different ways) been great sources of inspiration throughout my education. Louise taught me a lot about how to study when we lived together during my first years at university, which has given me the necessary work ethic to complete this thesis. Marie has been an inspiration as she has shown me that it is possible to keep fighting even if you lack oxygen.

Contents

Abstract **2**

1 Introduction **7**

2 Curricula **10**

3 Textbooks **13**

4 Didactical Preliminaries on Probability Theory **15**

4.1 Objectivist vs. Subjectivist 16
4.2 Misconceptions and Biases 21
4.2.1 Causal Reasoning and the Fallacy of the Time Axis 22
4.2.2 The Base Rate Fallacy 23
4.2.3 Fallacy of the Transposed Conditional 25
4.2.4 The Conjunction Fallacy 26
4.3 Intuition and Counter-Intuitive Problems 26
4.4 Linguistics and Context 28
4.5 Contingency Tables 29
4.6 Simulations 32

5 The Theory of Didatical Situations **35**

6 Design and A Priori Analysis **38**

6.0.1 General Information 39
6.1 First Exercise - A Contingency Table 40
6.1.1 TDS-Schedule 42
6.1.2 Analysis 43
6.2 Second Exercise - Another Contingency Table 47
6.2.1 TDS-Schedule 48
6.2.2 Analysis 49
6.3 Third Exercise - The Urn Problem 51
6.3.1 TDS-Schedule 52
6.3.2 Analysis 53
6.4 Fourth Exercise - The Monty Hall Problem 55
6.4.1 TDS-Schedule 55

6.4.2 Analysis 58
6.5 Fifth Exercise - The Drawer Problem 62
6.5.1 TDS-Schedule 63
6.5.2 Analysis 64
6.6 Sixth Exercise - The Three-Event Problem 67
6.6.1 TDS-Schedule 67
6.6.2 Analysis 68

7 Methodology 70

8 Results and A Posteriori Analysis 72
8.1 General observations 73
8.2 First Exercise - The First Contingency Table Problem 75
8.3 Second Exercise - The Second Contingency Table 80
8.4 Third Exercise - The Urn Problem 83
8.5 Fourth Exercise - The Monty Hall Problem 86
8.6 Fifth Exercise - The Drawer Problem 89
8.7 Sixth Exercise - The Three-Event Problem 92
8.8 Comprehension Tests 96
8.9 The Teachers and the Students' opinion 100

9 Discussion 101

10 Conclusion 104

References 106

A Appendix 110
A.1 Lesson Plans - Danish Version 110
A.1.1 First Exercise 110
A.1.2 Second Exercise 111
A.1.3 Third Exercise 112
A.1.4 Fourth Exercise 113
A.1.5 Fifth Exercise 116
A.1.6 Sixth Exercise 118
A.2 First Exercise - The First Contingency Table (Danish Version) . . . 120
A.3 Second Exercise - The Second Contingency Table (Danish Version) . 121

A.4 Third Exercise - The Urn Problem (Danish Version) 122
A.5 Fourth Exercise - The Monty Hall Problem (Danish Version) 123
A.6 Fifth Exercise - The Drawer Problem (Danish Version) 124
A.7 Sixth Exercise - The Three-Event Problem (Danish Version) 125
A.8 The Initial Comprehension Test 126
A.9 The Final Comprehension Test 127
A.10 The Initial Comprehension Test (Danish Version) 128
A.11 The Final Comprehension Test (Danish Version) 129

1 Introduction

Probability theory (especially conditional probability and independence) and the abilities and skills the students acquire to work with the subject are not only relevant for the students' general participation in society, but also within many different educations and future jobs (everything from political science, biology, history, economics to mason, banker, police officer, etc.). It is clear that academic work matters more in some professions than in others, where it is the mindset and skills that are relevant. But the central thing is that students' knowledge of conditional probability and independence is useful in their future, even to a greater extent than many other subjects in upper secondary school mathematics education. Unfortunately, probability theory is one of the subjects in mathematics teaching in upper secondary school that presents the greatest challenges. There is a large amount of research that shows that probability theory is the subject that upper secondary students find the most difficult and, furthermore, the most boring, (Batanero, Contreras, et al. 2014; Shaughnessy 1992; Carles and Huerta 2007, among others). For this reason, it is interesting, for the sake of society and the students, to investigate whether it is possible to improve teaching so that students have a better opportunity to acquire solid knowledge about probability theory (especially about conditional probability and independence). There are many possible reasons for the students' difficulty with probability theory, but one of the main reasons is, according to Batanero, Contreras, et al. 2014, that the teachers also think the subject is difficult. For me personally, this thesis is the culmination of my education as a mathematics teacher for upper secondary schools. I myself have considered how little the probability theory and statistics constitute of the education. When you educates to be a mathematics teacher in Denmark (for upper secondary school), it is most common that you only have one or two courses on probability and statistics, which corresponds to 7.5-15 ECTS points (that is 2.5-5% of the education), which is little compared to most other branches of mathematics.

In addition to not having received much education in probability theory and statistics, I have always found the subject either difficult and uninteresting or easy and boring. I often thought we wasted time in pre-university teaching when we had to roll 100 times with a die or perform 100 coin tosses, only to find that the die or coin behaved as we expected. After enduring this monotonous, exhausting and repetitive exercise, one often had to rely on strange problems that were incomprehensible. These problems seemed set up and it was often difficult to find out what the problem actually was and what to do to solve it. These difficulties in understanding the problems are related to the fact that probability theory and statistics have different terminology and thus language is used much more than in other branches of mathematics. Not only is the terminology different, it is also "*clumsy*", as it was called by Feller 1968, and this is especially true of conditional probability and independence.

According to Borovcnik and Peard 1996, there are particularly many counterintuitive problems within probability theory, which probably has a contributing force in the fact that the subject seems difficult and inaccessible to many students and teachers. Borovcnik and Peard 1996 wrote that in addition to having several counterintuitive problems, these problems spread over all levels. In other branches of mathematics, it is more normal that counterintuitive problems first arise at a higher level of abstraction, but this is not the case with probability theory. Several researchers, Shaughnessy 1992; Batanero and Sanchez 2005; Díaz and Batanero 2009, have shown that there are a large amount of misconceptions within probability theory, but especially within conditional probability and independence. These misconceptions often lead to a false reasoning for answers to problems within the subject, especially in the case of one of the many counterintuitive problems.

For all these reasons, it is of particular interest to explore the possibilities of optimizing the teaching of conditional probability and independence, so that fewer students are left with misconceptions that are not conducive to their future career and social participation. To frame my research, I have created the following main research questions:

> *How should conditional probability and independence be taught so that students acquire solid knowledge about the subject that is conceptualized in a non-misleading way?*

To clarify this question, there are three smaller research questions below, all of which aim to answer parts of the main question

- What misconceptions about conditional probability and independence are widespread among upper secondary school students and is it possible to avoid them?
- How can counterintuitive problems be used in the teaching of conditional probability and independence and what effect does this have on students' mathematical output and conceptualisation?
- How are students' linguistic and formalization abilities strengthened within the terminology of probability theory through a course of around 4 hours?

In order to investigate these questions, I will explain how conditional probability and independence are taught to students in upper secondary school today and within the last 15-20 years. The study of this will be based on curricula and teaching materials, and a comparison will be made with how the subject is taught in other countries with comparable conditions. The didactic research in the area will then be explained, where the focus will be on which misconceptions occur most often and how to avoid them. There will be a brief explanation of the Theory of Didactic Situations, which has been used as a framework for the teaching design I have developed. The teaching design has been developed based on a didactical a priori analysis, made on the basis of the research that has been explained. The teaching design has been tested by an upper secondary school class, which will be kept anonymous. A great thanks shall go to the school management for allowing me to try out my teaching design in their school, but the greatest thanks goes to the students of the class who all participated eagerly in the teaching situations and approached the challenges with an open mind, and to the class teacher, who with great effort took on the difficult task of teaching based on a design he did not make himself, and who also was willing to share his thoughts on the process. The trial led to the collection of a large amount of data which has been analysed. The results of this analysis will be presented at the end of the thesis.

2 Curricula

In the history of Danish mathematics curriculum, there have been three curricula since 2005, when there was a major reform on the upper secondary school area. The current curriculum for mathematics at the highest level in the Danish upper secondary school (called A-level) is from 2017, and is referred to as *Læreplan Matematik A - STX* 2017, and before that a curriculum in 2013 (*Læreplan Matematik A - STX* 2013) and in 2005 (*Læreplan Matematik A - STX* 2005). In the current curriculum, there is nothing about conditional probability. In the section "Kernestof og mindstekrav", which can be translated to "Core materials and minimum requirements", *Læreplan Matematik A - STX* 2017 says the following about probability theory

> The core materials are combinatorics, basic probability calculus, probability field and random variable, binomial distribution and normal distribution, confidence intervals, hypothesis testing in the binomial distribution. [Translated into English by the author] *Læreplan Matematik A - STX* 2017, p. 2.

It is clear that basic probability calculus is not firmly defined, which means that conditional probability and independence could well be part of it. Especially when the binomial distribution is actually on the list of core materials, since independence is a prerequisite for the binomial distribution. This curriculum has a guide attached, which elaborates and explains the syllabus in greater detail. In this guide the following is stated in contrast to the above rationale

> Independent events are mentioned in connection with problem solving that requires multiplication of probabilities, but are not given a standalone treatment. [Translated into English by the author] *Vejledning Matematik A - STX 2017* 2021, p. 8.

Since the guide is also defining for the syllabus, this means that independence is not a part of the syllabus. The same applies to random variables, since in the guide (*Vejledning Matematik A - STX 2017* 2021) there is a sentence that is almost identical to the quote above, only with random variables instead of independence. This is noticeable, as the following is stated later in the guide.

> Students must know the conditions for when an empirical data set can be considered as actual values of a binomially distributed random variable; including a discussion of experiments with and without replacement. [Translated into English by the author] *Vejledning Matematik A - STX 2017* 2021, p. 9

This is noticeable, since random variables and independence is fundamental for binomial distribution. It is very hard to see, in which way students can know the conditions of a binomially distributed random variable without knowing anything about random variables and independence.

The two previous curricula are both slightly different on this point, but in their own ways. Back in 2005, the curriculum was characterized by not including probability theory in the core material, despite statistics being included. Probability theory was included only in the optional extension material. This was criticized by many, including EVA, which is the Danish evaluation institute, in EVA 2009. Someone may notice that the criticism from EVA only came in 2009, i.e. a full four years after the reform in 2005, which introduced the curriculum *Læreplan Matematik A - STX* 2005. This is because it takes three years for students to study an A-level, so EVA waited to make their evaluation of the curriculum until after there has been a whole year of students through the A-level in the upper secondary school. The criticism is that EVA's experts find probability theory important for the student's understanding of the society they live in. So the experts focus more on the cultural and social aspects of the theory than on the actual application. The curriculum's argument for moving probability theory from the core material to the extension material was that it became easier for the teachers to select and put together the teaching plan based on what made the most sense for the class and the study package of the class. EVA's experts recognize that statistics and probability theory are areas of mathematical teaching that interacts well with other subjects, for instance social studies and biology are two subjects that really could benefit from an interaction with statistics and probability. However, Eva's experts believe that probability theory is fundamental for teaching statistics, and they strongly wonder why statistics is in the core material when probability theory is left out.

After the criticism of the curriculum from 2005, a new curriculum was created in 2013 (*Læreplan Matematik A - STX* 2013), but despite the criticism, probability theory was still not part of the core material. In return, independence was included in this curriculum since test of independence were now a part of the syllabus, in the form of chi-square test. The reason the chi-squared test was included in the curriculum was that this test of independence was used in the subjects of biology and social studies as a statistical tool. In today's curriculum, the chi-squared test is no longer included, but it is still included in both biology and social studies curriculum, as a statistical tool, meaning that students are taught chi-squared test by either their biology teacher or their social studies teacher. Since these teachers are most often not mathematics teachers, too, and thus not trained to teach mathematics, this construction hardly helps the students to better conceptualize probability or statistics.

If we look outside the borders of Denmark, to countries that are in many ways comparable to Denmark in the school and teaching areas on the basis of culture, economy and quality of life, we see a slightly different content in curricula. Let us start with the US. There is no set curriculum for the entire United States, as each state determines the curriculum by themselves. However, in 2010 there was a common standard, called *Common Core State Standards*, which was developed by two non-profit associations, which has formed a foundation for curricula in the United States. As many as 40 states chose to use Common Core as their curriculum or at least as the basis of their curriculum, which makes it relatively representative of curricula in the United States (Molnar 2015). So if we look in *Common Core State Standards for Mathematics* 2022, we first of all see that conditional probability and independence are both part of the curriculum. The Common Core has nine points that describes what must be reviewed in the teaching of conditional probability and independence. The nine points are very reminiscent of how the mathematical theory was reviewed in this thesis in section 2. This means that the American curriculum is far more comprehensive within conditional probability and independence than the Danish one.

In England, conditional probability and independence are already part of the curriculum for what they call *key stage 4*, which corresponds to the 14-16 age groups (*The National Curriculum of England* 2014). The approach is different from the American one, as the English introduce conditional probability through the more illustrative representation using expected frequencies with contingency tables, tree diagrams and Venn diagrams.

In Denmark, it is quite normal to compare ourselves with our Scandinavian neighbours, and even countries that are so close to Denmark stand out from the Danish curriculum. In Sweden, all upper secondary students learn about independence, no matter what level of mathematics they study (*Curriculum of Sweden* 2022). This priority says something about how important independence is assessed by the Swedish Board of Education (*Skolverket*). In addition, correlation and causality are also both in the curriculum. It will be explained later in this thesis that correlation and causality is a good support for independence and conditional probability, and strengthens the connection between statistics and probability.

In Norway, they have a bit more complicated school system for the upper secondary level. Students can choose for themselves whether they want to take mathematics courses that are either theoretical or practical at the first year. After that, they can (very simplistically) choose between continuing a theoretical path (if they had chosen this path in their first year), or a social studies path. What is interesting, however, is that students who choose the theoretical path will never encounter probability theory nor statistics. Those students who choose the social science path will encounter little statistics and no probability theory. Conditional probability and independence are not a part of the curriculum in Norway (*Curriculum of Norway* 2022).

3 Textbooks

As has just been explained in the previous section, conditional probability and independence are not part of the curriculum in Denmark and have not been since 2017. Therefore, conditional probability and independence are included only to a very limited extend. Typically, independence is only included as a form of minor notice in connection with the binomial distribution, and conditional probability is not mentioned in any of the textbooks reviewed in connection with this thesis. For this reason, it is more interesting to look at the textbooks that were made for the period where *Læreplan Matematik A - STX* 2013 was applicable.

From 2005 until 2013, neither conditional probability nor independence were part of the syllabus, as previously mentioned. However, it is extremely interesting that both parts were included in several textbooks from this period. If we start by looking at the textbook system from the publisher *Gyldendal*, which consists of a basic book with the basic theory, Clausen et al. 2006b, and an exercise book with exercises and a bit of extra theory, Clausen et al. 2006a. Conditional probability and independence are only included in the exercise book, except that it is mentioned that the experiments in a binomial distribution must be independent, which is very similar to contemporary textbooks. This is probably because it was possible to bring up the subject as an additional subject. The exercise book goes through conditional probability and independence in quite a lot of detail. The exercise book contains a section with an experimental approach to binomial distribution. As a basis for this section, Clausen et al. 2006a have created an introductory section called "*Rules for calculating probabilities*" (translated by the author), which, with the help of e.g. Venn diagrams, the notation from set theory and examples, define both conditional probability, Bayes' theorem for two events and in the general form, independence for two events and the multiplication principle for two events and then in the general form.

One of the largest school book publishers in Denmark is the publisher *Systime*, which published the mathematics book Carstensen et al. 2006 in connection with the new curriculum from 2005. In this textbook there is also a section on conditional probability and independence, the section is simple labeled "Independence" (translated by the author), which, just like in the other book, is a precursor to binomial probabilities. However, unlike Clausen et al. 2006a, Carstensen et al. 2006 use contingency tables to form a basis for the definition of conditional probability, then define independence using conditional probability (as in definition 2.6), and end with the multiplication principle as a theorem without proof.

Systime has produced several textbooks for upper secondary mathematical teaching, e.g. they published Brydensholt and Ebbesen 2011, where they had gained a lot of experience with the curriculum and its possible challenges. Unlike Carstensen et al. 2006, which is also a book from the publisher Systime, this book starts with independence and waits with conditional probability for the section after. For this reason, Brydensholt and Ebbesen 2011 also had to define independence. This definition is different from the definition in the other textbooks, as this book uses the multiplication principle as the definition, unlike the other books that make the definition based on conditional probability. This textbook differs even more as it defines independence for three events. The textbook goes through conditional probability with the same procedure as the other textbooks, however this textbook uses tree diagrams and a lot of examples to elaborate the theory.

As stated in the previous section about curricula, the chi-squared test was in extension material in *Læreplan Matematik A - STX* 2005, but was moved into the core material, which is the syllabus, in *Læreplan Matematik A - STX* 2013, when students had to use the test in other subjects, such as social studies and biology. In *Læreplan Matematik A - STX* 2017 the chi-squared test disappeared completely from the curriculum, so it does not appear in new textbooks, but it was in the vast majority of textbooks from 2005 to 2017. The chi-squared test can be both a hypothesis test and a test for independence. In the textbooks, it is most often used in connection with contingency tables. You can read more about contingency tables in section 4.5.

4 Didactical Preliminaries on Probability Theory

As previously mentioned, it is widely described in the literature that conditional probability and independence is an area that has certain challenges in terms of being taught by teachers and conceptualized by students. Historically speaking, probability theory and statistics first appeared relatively late (Batanero, Henry, et al. 2005). According to Székely 1986, there are many indications that the Arabs invented the dice, and thereby started to work a little with simple probabilities. Gerolamo Cardano (1501-1576), Galileo Galilei (1564-1642) and other Renaissance mathematicians did the same, but without the further abstraction. In 1654 Chervalier de Méré (1607-1684) contacted Blaise Pascal (1623-1662) to seek help in solving a problem of how to fairly divide the stakes in an interrupted game. Pascal and Pierre de Fermat (1607-1665) solved this problem, which today is called the problem of points, independently. Precisely the incident is described by many as the start of the theory of probability, as Pascal used methods that are today used in probability theory. The reason this is interesting is that this historical hesitation suggests a requirement for a relatively high level of abstraction in order to work with probability theory.

Through this chapter, it will be reviewed which misconceptions and beliefs are contributing to giving probability theory this high level of abstraction. This is important in relation to the later work of developing a design for teaching conditional probability and independence, as it provides prior knowledge of the students' lack of understanding of the subject.

4.1 Objectivist vs. Subjectivist

In science, there is usually a requirement for objectivity. This requirement leads to certain problems within probability theory and statistics. Therefore, there are two different approaches to the requirement for objectivity within these subjects. There is the objectivist and the subjectivist approach. There are those who believe that objectivity must of course be maintained no matter what. They are called objectivists. One of the central objectivists is Andrey N. Kolmogorov. He begins his work *Foundations of the Theory of Probability* with a sentence that describes an objectivist's approach to probability.

> "*The theory of probability, as a mathematical discipline, can and should be developed from axioms in exactly the same way as Geometry and Algebra.*" Kolmogorov 1933/1956, pp. 1

Therefore, the objectivist believes that probability must be separated from personal judgments. Borovcnik 2012 explains how, for an objectivist, probability is a property inherent to an object. Kolmogorov created an axiomatic theory of probability that can be used to indirectly determine probabilities, where probabilities are considered a property of an object. The following are Kolmogorov's axioms for probability theory, but with a slight difference in notation, as a uniform notation is desired throughout this thesis.

> Let S be a collection of elements $\varepsilon, \eta, \zeta, ...$, which we shall call *elementary events*, and T a set of subsets of S; the elements of the set T will be called *random events*.
>
> I T is a field of sets.
>
> II T contains the set S.
>
> III To each set A in T is assigned a non-negative real number $P(A)$. This number $P(A)$ is called the probability of the event A.
>
> IV $P(S)$ equals 1
>
> V If A and B have no element in common, then
>
> $$P(A \cup B) = P(A) + P(B)$$
>
> Kolmogorov 1933/1956, p. 2

From these axioms, a number of properties of probabilities can be determined. Here are three key properties, theorem and proof inspired by Blitzstein and Hwang 2014

Theorem 4.1. *Probability has the following properties, for any events A and B.*

a $P(A^c) = 1 - P(A)$.

b *If* $A \subseteq B$, *then* $P(A) \leq P(B)$.

c $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Proof. a By the construction of the complement, it follows that $A \cup A^c = S$, and that A and A^c are disjoint. According to Kolmogorov's axiom IV, this implies

$$1 = P(S) = P(A \cup A^c) = P(A) + P(A^c)$$

If $P(A^c)$ is subtracted on both sides, we have

$$P(A) = 1 - P(A^c)$$

b Suppose that $A \subseteq B$, then $B = A \cup (B \cap A^c)$. It is clear, by the construction of the complement, that A and $(B \cap A^c)$ are disjoint, and then Kolmogorov's axiom V is useful

$$P(B) = P(A \cup (B \cap A^c)) = P(A) + P(B \cap A^c)$$

Since $P(B \cap A^c)$ is a probability, and a probability is non-negative (axiom III), the following is true

$$P(B) \geq P(A)$$

c With the same argument as above, it holds that $A \cup B = A \cup (B \cap A^c)$ and that A and $B \cap A^c$ are disjoint. By Kolmogorov's axiom V, the following holds

$$P(A \cup B) = P(A \cup (B \cap A^c)) = P(A) + P(B \cap A^c)$$

Likewise, $B = (A \cap B) \cup (B \cap A^c)$ where $A \cap B$ and $B \cap A^c$ are disjoint, and the following holds

$$P(B) = P((A \cap B) \cup (B \cap A^c)) = P(A \cap B) + P(B \cap A^c)$$

Isolating $P(B \cap A^c)$ and substitute into the previous equation:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

□

For an objectivist, it is important to be able to conceptualize the repetition of a random experiment, which is done using the concept of independence. In order to define independence, conditional probability have to be defined. This definition is inspired by Sørensen 2013, p. 22:

Definition 4.1. *The conditional probability* of the event A given another event B (with non-zero probability, $P(B) > 0$), written $P(A|B)$, is defined by

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

This concept is justified by fulfilling Kolmogorov's axioms, which is proven here:

Theorem 4.2. *Conditional probabilities are probabilities*

Proof. To check whether conditional probabilities are probabilities, it is enough to check whether the last two of Kolmogorov's axioms are satisfied.

IV It is trivial that $P(S|E) = 1$ holds, where S is the sample space and E is an event in the sample space, since the probability that the entire sample space happens is unaffected (or unconditional) by the event E.

V If it is supposed that $A_1, A_2, ...$ are disjoint, and we look at $P(\bigcup_{j=1}^{\infty} A_j|E)$ using definition 4.1:

$$P\left(\bigcup_{j=1}^{\infty} A_j|E\right) = \frac{P\left(\left(\bigcup_{j=1}^{\infty} A_j\right) \cap E\right)}{P(E)} = \frac{P\left(\bigcup_{j=1}^{\infty}\left(A_j \cap E\right)\right)}{P(E)}$$

The last equality is true according to normal set theory. Since the numerator of the fraction is a probability axiom V holds, and the following must hold

$$P\left(\bigcup_{j=1}^{\infty} A_j|E\right) = \frac{P\left(\bigcup_{j=1}^{\infty}\left(A_j \cap E\right)\right)}{P(E)} = \frac{\sum_{j=1}^{\infty} P(A_j \cap E)}{P(E)}$$
$$= \sum_{j=1}^{\infty} \frac{P(A_j \cap E)}{P(E)} = \sum_{j=1}^{\infty} P(A_j|E)$$

The final step is once again by using definition 5.1.

□

With the definition of conditional probabilities, it is now possible to define independence.

Definition 4.2. Two events with positive probabilities, A and B, are said to be *independent* if

$$P(A|B) = P(A)$$

or equivalently

$$P(B|A) = P(B)$$

Independence is a central concept for the objectivist, as it is the key assumption in several central theorems, as the law of large numbers and the central limit theorem. It is so central that Borovcnik writes the following:

> *For an objectivist conception of probability, independence is key and crux.*
> Borovcnik 2012, pp.8

Besides the main point that all science must be objective, including probability theory, one of the rationales behind the objectivist approach is that it can help escape the trap of intuition. There are numerous examples of counterintuitive problems and paradoxes within probability theory, e.g. The Monty Hall problem, the best-known counterintuitive problem from probability theory. Therefore, to a much greater extent than in other mathematical fields, one must be aware that one's intuition can be misleading. The easiest and best tool to avoid this intuitive delusion trap is to be objective, and thus not relate to one's intuition.

The objectivist approach has several points of criticism. The first and most central is the critique that the concept of independence is central to the structure of probability theory. A critic will say that independence is an idealization, and thus the theory is largely impossible to use in practice. Another point of criticism is that the objectivist does not include statistical inference. To include this it would require additional terms and criteria that are not acceptable to an objectivist, as it would require an acceptance of irreparable gaps in the logical basis or admit of a subjectivist nature of probability theory. Last but not least, the key concept, independence, according to Borovcnik 2012, lacks a link to mental images and reference situations. This means that probability, conditional probability and independence become "*a permanent source of confusion* (Borovcnik 2012, pp. 8).

In contrast to the objectivist, there are also those who believe that it is not possible to be objective in probability theory and statistics. They are called subjectivists. This approach is closer to what is called provability, which involves finding more arguments for a statement than arguments against it, than the objectivist approach. Thus, objectivity is replaced by an judgment of credibility. There are different types of credibility, e.g. expert knowledge or personal expectations. There is one and only one credibility that is also accepted by the objectivist, namely relative frequencies from relevant experiments in the past. This is accepted by the objectivist, as these relative frequencies can be assessed without subjective judgments. The subjectivist approach is, like the objectivist, justified axiomatically. The axioms that can be found in Finetti 1974, unlike in the objectivist position, are based on rational behavior and describe criteria for our preferences. Or as Borovcnik puts it, "*the paradigmatic situation is betting on uncertain statements*" (Borovcnik 2012, pp. 9). According to a subjectivist, their approach is as valid as the objective one, since the axiomatic foundations (Kolmogorov 1933/1956 and Finetti 1974) have the same scientific status. Furthermore, it is fundamental to a subjectivist's argument that if the amount of empirical data becomes sufficiently large, their conclusions will be the same as the objectivist's conclusion. Subjectivists believe that as the amount of information increases, the judgment will also become more qualified, and thus the probability will become more authorized. The largest critique of the subjectivist is of course the lack of objectivity, but especially the lack of empirical control of random experiments is criticised. It is usually a crucial requirement for empirically based science that there is a control on the empirical evidence. As central as independence is to the objectivist, so is conditional probability to the subjectivist. A subjectivist believes that all probabilities are conditional, and should be updated as new evidence comes in, in order to get a more and more accurate result. To do this kind of updates of probabilities, Bayes' theorem is useful. This means that Bayes' theorem plays a key role in the subjectivist position. For the objectivist, Bayes' theorem is not nearly as interesting, as it is a consequence of Kolmogorov's axioms.

Theorem 4.3 (Bayes' theorem)**.** *For any event A and B with positive probabilities, the following holds*

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Proof. From definition 4.1 it is known that

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \qquad \text{and} \qquad P(B|A) = \frac{P(B \cap A)}{P(A)}$$

Since $A \cap B = B \cap A$ the following holds

$$P(A \cap B) = P(B)P(A|B) = P(A)P(B|A)$$

Dividing by $P(B)$ on both sides of the last equality gives

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

□

As a remark to this theorem, the multiplication principle can be proved:

Theorem 4.4 (The Multiplication Principle)**.** *For two independent events, with non-zero probability, the following holds*

$$P(A \cap B) = P(A)P(B)$$

Proof. From definition 4.1 it is known that

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \Rightarrow P(A \cap B) = P(A|B) \cdot P(B)$$

$P(A \cap B) = P(B|A) \cdot P(A)$ could be used in parallel with this proof. Since A and B are independent it holds that $P(A|B) = P(A)$, according to definition 4.2, which means that

$$P(A \cap B) = P(A)P(B)$$

□

This theorem is equivalent to definition 5.2, and is therefore often used as the definition for independence.

4.2 Misconceptions and Biases

It has just been established that there are more examples of counterintuitive problems in probability theory than in other mathematical fields (spread out at all levels). This large amount of counterintuitive problems originates from the fact that there is a wide spread of misconceptions and biases among both students and teachers (Díaz and Batanero 2009, Batanero, Contreras, et al. 2014). Below, a brief description of selected misconceptions and biases will be given, which will be elaborated with examples of problems where the relevant misconception or bias could mislead a person who tries to solve the problem. We start with one of the significant misconceptions that have several branches of misconception.

4.2.1 Causal Reasoning and the Fallacy of the Time Axis

Gras and Totohasina 1995 identified three different misconceptions about conditional probability via their empirical data from questionnaires of 17-18 year old secondary school students. The three misconceptions were; the chronological conception, the causal conception and the cardinal conception. The chronological conception is that students believe that an event cannot condition another event that happens before it (Falk 1989). Which means that students think that $P(A|B)$ is meaningless if event B occurs after event A, since these students see conditional probability ($P(A|B)$) as a temporal relationship where event B always precedes event A. This is also called the time axis fallacy. The causal conception is that students believe that $P(A|B)$ means that event B is the cause and event A is the consequence. Thus, it is the causality and not the timeline that is the conceptual obstacle for the student. These two concepts are clearly closely related, since a causal event will naturally precede a consequential event. So the chronological conception can be considered as a special case of the causal conception. The reason why the time axis fallacy and the causal fallacy are not considered as a single misconception is because people who make these fallacies can be divided into those who justify it with a time axis argument and those who justify it with a causality argument. The cardinal conception is an offshoot of a general misconception in probability theory, namely that there is equal probability in every probability space. That is that students who have this fallacy will think that $P(A|B) = \frac{|A \cap B|}{|B|}$, where $|A \cap B|$ is the cardinality of the set $A \cap B$, and $|B|$ is the cardinality of the set B. Precisely this misconception is not related to causal reasoning, and will not take up more space in this thesis.

A classic problem that illustrates the problem of causal reasoning or the time axis fallacy is this problem:

> An urn contains two white balls and two red balls. We pick up two balls at random, one after the other without replacement. (a) What is the probability that the second ball is red, given that the first ball is also red? (b). What is the probability that the first ball is red, given that the second ball is also red?
> (Batanero and Sanchez 2005, p. 251)

In this problem, part *a* will not cause much difficulty, according to Batanero and Sanchez 2005 and Falk 1989. Since the condition that a red ball was drawn in the first draw will narrow the sample space so that there are only two white balls and one red ball, all with equal probability, the conditioned probability is $P(B|A) = \frac{1}{3}$, where A is the event to get a red ball in the first draw and B is the event to get a red ball in the second draw. People affected by the above mentioned misconceptions will not see the problem differently than people that are not affected by the misconceptions, since the condition does not work against the direction of time, or since the first move clearly causes a narrowing of the sample space. For part *b* of the problem, according to Batanero and Sanchez 2005 and Falk 1989, there will be many who will give a wrong answer by false reasoning, namely $\frac{1}{2}$. The solution is parallel to the solution to part *a*, since the information that a red ball has been drawn in the second move narrows down the sample space for the event A in exactly the same way as before, so that the probability again becomes $P(A|B) = \frac{1}{3}$. People with one of the misconceptions will, as mentioned, give the answer $\frac{1}{2}$. Due to one of the fallacies they consider the conditional probability $P(A|B)$ to be meaningless, instead they determine the unconditional probability $P(A) = \frac{1}{2}$.

In general, there is a tendency for people to overestimate causality and its importance when working with probability, where a diagnostic approach would be more accurate. This is probably because the impact of causal data on a judgment is usually greater than the impact of diagnostic data. In addition, it is quite normal for human beings to relate to causes, while diagnostics are more advanced for us. Human beings are, so to speak, more used to causes than to diagnostics.

4.2.2 The Base Rate Fallacy

The base rate fallacy was identified by Tversky and Kahneman 1982 and is best illustrated with an example:

> Why is more grass consumed by white sheep than by black sheep?
> (Bar-Hillel 1983, p. 39)

The fallacy is that people think that more of the grass is consumed by white sheep because white sheep consumes more grass than black sheep and thus forget the base rate, namely that there are far more white sheep than there are black sheep. There are unimaginable amounts of this type of example, many originate from everyday life, where e.g. news media may come up with stories based on this fallacy. In connection with schoolwork, it is often desired that the problems posed lead to some form of calculation. These types of problems are often Bayesian, but the fallacy remains that people forget to include the base rate (or prior probability). Bayesian problems are problems that can be solved using Bayes' theorem (Theorem 4.3). Normally, there is two types of information, which are needed to solve a Bayesian problem. A generic information, which is about the frequencies of the hypothesis, this information is called the base rate. In addition, specific information have to be used to deal with the case in the Bayesian problem. Some people tend to base their answer to a Bayesian problem solely on the specific information, which is a base rate fallacy.

A classic example illustrating how this fallacy can play into a Bayesian problem could be

> A witness sees a crime involving a taxi in a city. The witness says that the taxi is blue. It is known from previous research that witnesses are correct 80% of the time when making such statements. The police also know that 15% of the taxis in the city are blue, the other 85% being green. What is the probability that a blue taxi was involved in the crime?
> (Díaz and Batanero 2009, p. 157)

A person could be led to believe that the probability that there was a blue taxi involved in the crime must be 80% since the witness is telling the truth 80% of the time when they make such statements. It would then be a base rate fallacy, as the person forgets that there are only 15% blue taxis in the city. In order to solve this problem, Bayes' theorem is useful as mentioned.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Let A be the event that the taxi involved in the crime was blue, and B be the event that the witness says that the taxi was blue. Thus, the following information is given in the problem: $P(B|A) = 80\%$ and $P(A) = 15\%$. Unfortunately, $P(B)$, i.e. the probability that the witness says that the taxi was blue (when it is unknown whether the taxi was blue or green), is not stated. Therefore, it is necessary to use Bayes' theorem in an alternative form:

Theorem 4.5 (Bayes' Theorem [Alternative Form]). *For any events A and B with positive probabilities, the following holds:*

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}$$

Proof. The normal form of Bayes' theorem looks like this

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Thus, it is enough to show that $P(B) = P(B|A)P(A) + P(B|A^c)P(A^c)$. Classical set theory says that $B = (B \cap A) \cup (B \cap A^c)$. Since $(B \cap A)$ and $(B \cap A^c)$ are disjoint, it must hold that

$$P(B) = P\Big((B \cap A) \cup (B \cap A^c)\Big) = P(B \cap A) + P(B \cap A^c)$$

From the definition of conditional probability (definition 4.1) it is seen that $P(B \cap A) = P(B|A)P(A)$ and $P(B \cap A^c) = P(B|A^c)P(A^c)$, which thus means

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}$$

$\square$

It is clear that $P(A^c) = 85\%$ as it is the probability that the taxi is not blue (ie. green in this case). $P(B|A^c)$ is the probability that the witness says that the taxi was blue, when it was actually green. Since it is known that there is 80% chance that the witness is telling the truth when she says that the taxi was blue, it holds that $P(B|A^c) = 1 - 0.80 = 20\%$, and only the calculation is left.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)} = \frac{0.80 \cdot 0.15}{0.80 \cdot 0.15 + 0.20 \cdot 0.85} \approx 0.41 = 41\%$$

4.2.3 Fallacy of the Transposed Conditional

Bar-Hillel 1983 believes that in the taxi problem (from the previous section) and in problems of the same type, there is a possibility that another fallacy can arise, namely the fallacy of the transposed conditional. This fallacy occurs when people cannot tell the difference between $P(A|B)$ and $P(B|A)$. In the taxi problem, it is essential to distinguish between $P(A|B)$, which is the probability that the taxi was blue, given that a witness says that the taxi was blue, and $P(B|A)$, which is the probability for the witness to say that the taxi was blue, given that the taxi actually was blue. In the taxi problem, it can be difficult to distinguish between this fallacy and the base rate fallacy, since the fallacies are very close to each other, and since they both will lead to the same incorrect conclusion. The fallacy of the transposed conditional can arise in problems where the base rate fallacy is not relevant, as it is possible to exchange the conditioned and the conditioning event in most types of problems within conditional probability, while the base rate fallacy of course requires a problem with a base rate.

4.2.4 The Conjunction Fallacy

This fallacy is about how people misestimate the probability of a conjunction. Often it can be seen by a person assessing the probability of A and B to be greater than the probability of A (or B), i.e. $P(A \cap B) > P(A)$. This is not true, since it is always true that $A \cap B$ is a subset of A (and B), which is known from normal set theory. Therefore, it holds that $P(A \cap B) \leq P(A)$. An example of a situation where this fallacy could occur is

> (a) What is the probability that a random employee in the public sector has problems with pain in the back? (b) What is the probability that a random employee in the public sector is 55 years old and has problems with pain in the back?

Here the conjunction fallacy will be that there is a greater probability that the random person is 55 years old and has problems with back pain than that the random person simply has problems with pain in the back. The fallacy is due to an exchange of conditional probability and a conjunctional probability. So instead of determining the probability that the random person is 55 years old and has back pain problems, they determine the probability that a random 55 year old has back pain problems, so $P(A|B)$ instead of $P(A \cap B)$, where A is the event of having problems with back pain and B is the event of being 55 years old. The fallacy can also be an underestimation instead of an overestimation as here.

4.3 Intuition and Counter-Intuitive Problems

All human beings have used the same intuitive method, since we were very young. We encounter problems that we have not been exposed to before, and we therefore do not know how to tackle. We try what we believe in most, and if it does not work, we adjust and try again. Someone will recognize this Deweyian method as "*trial and error*. Within mathematics, it is also a widespread and quite successful method, no matter whether it is for academic research or for school students' method of solving exercise problems. Unfortunately, it is the case that trial and error does not work nearly as well within probability theory as within other mathematical fields. If a person tries to predict an event, the person can easily be right, even if the person predicted something that was less likely than something else. It could, for example, be a child who has to roll two dice and try to predict the sum of the outcomes of the two dice. The child knows nothing about probability and naively guesses that the sum will be four. Then the child rolls and the sum of the two dice happens to be four. Thus, the child incorrectly thinks that four was the best choice. However, it is not only children who, with their naive approach to probability theory, come to make fallacies based on trial and error. It can also happen to adults, of which the whole idea of astrology is an excellent example.

In probability theory, concrete operations are missing, according to Borovcnik and Peard 1996, which leads to an immediate lack of solution tools, which makes it more attractive to approach the problem with intuition. Furthermore, it is the case that in probability theory, according to Borovcnik and Peard 1996, counterintuitive results occur at all levels of abstraction, in contrast to how it is in most other mathematical areas, where the counterintuitive results only occur at a higher level of abstraction. The combination of the desire to go to problems with intuition and the amount of counterintuitive results already at relatively low levels of abstraction often leads to learning difficulties and to some people having difficulty accepting the conceptions of probability theory. However, this does not mean that counterintuitive problems should be avoided, as students should train their abilities to deal with this type of problem, as they are so widespread as they are. In addition, it is particularly satisfying and enlightening, when it is possible to defy intuition with reflected and theoretical arguments. Working with counterintuitive problems often forces the person that is trying to solve the problem, to bring theory into play, as opposed to non-counterintuitive problems, where the person could give a non-theoretical but subjective and intuitive explanation.

The most important thing for learning probability theory is concept formation. In teaching, the discipline; that students must form concepts, is unfortunately difficult, as it cannot be provoked by a hierarchical sequence of actions and reflections. The way to concept formation is through the students' intuition, although it has just been described as being a cause of learning difficulties. But, according to Borovcnik and Peard 1996, theoretical inputs help students to revise their intuitions, which are basically raw and primitive, but which with each theoretical input will develop into a more and more complex and strong conception about probability theory. The theoretical inputs are important in this process, as the students will otherwise rely on a subjective approach that will be dominated by idiosyncratic and uncontrollable intuitive thinking, as there is no direct feedback and control of their subjective experiences. The theoretical input helps students towards a more objective approach (which will not be fully objective, however). This learning process or concept formation can be compared to how one generally develops new theoretical mathematics at an academic level.

4.4 Linguistics and Context

In all areas of mathematics, linguistics is important, but precisely within probability theory the importance is at its peak. When a teacher has to create a problem for her students, it requires great care in the formulation of the problem, just to make the problem comprehensible and without ambiguous wording. When the students have to work with the problem, it requires at least as much care for the students when they have to understand the problem and often requires several readings of the wording of the problem. This means that many problems within probability theory have two very different challenges that must be overcome, namely the purely mathematical challenge and the purely linguistic challenge. Often (especially at the level students are in upper secondary school) the linguistic challenge is almost the hardest part of the problem, which becomes even more prevalent when talking about conditional probability. Already in 1968, the probability theorist William Feller expressed the following quote, which describes exactly how the linguistics or terminology is challenging when working with probability theory, especially with conditional probability.

> "*The notion of conditional probability is a basic tool of probability theory, and it is unfortunate that its great simplicity is somewhat obscured by a singularly clumsy terminology.*" (Feller 1968, p. 114)

As mentioned, this means that there is a lot of focus on the language when working with probability, especially in teaching. For this reason, Shaughnessy 1992 suggested that the problems students had to work on in connection with teaching should be "*context free*". It is debatable whether it is possible to make a context free problem or whether there always will be a sort of context. In any case, it is clear that there are problems where the context is as basic and sterilized as possible (e.g. problems with rolling dice, tossing coins, urns with colored balls, draw a card from a deck etc.). The opposite of that is problems where the context looks like (or at least tries to look like) an everyday setting. Obviously, a sterilized problem makes it easier for students to focus on the specific problem. On the other hand, the everyday problems are more relevant for the students' future use of probability theory, at least how they should approach probability problems. Thus, a mixture of problems of the two types of context can be an advantage. Students find it easier to understand the nature of the problem in the sterilized version, and can use these problems as a reference point when solving problems in an everyday context. In conditional probability problems, students often have difficulty defining what there is a conditional probability. Especially in problems that are not sterilized, students can almost drown in what they feel is information, causing them to have difficulty analyzing the problem. Therefore, it ends up with the students determining a different object than the intended one. The difficulty of defining conditional probabilities is also seen with sterilized problems. Therefore, it requires a lot of training and practice in the use of the clumsy terminology before the students are able to analyze the problem correctly.

It was mentioned earlier that there are two equivalent definitions of independence; one that uses conditional probability ($P(A|B) = P(A)$), and one that is often used as a property of independent events and which is sometimes called for the multiplication principle ($P(A \cap B) = P(A)P(B)$). It turns out that it is advantageous to introduce independence using the conditional probability definition rather than the definition also called the multiplication principle. According to Kelly and Zwiers 1988, it is easier for students to understand the conditional probability definition, as it is close to how most people define the word ”independence” linguistically. A frequently used definition of the word is indeed ”something that is not affected by anything else”, which is aligned with the mathematical definition. Kelly and Zwiers 1988 argue that language matters a lot in probability theory (which has just been explained), and therefore, it is important that the definition is close to how the language is used, because then the students' intuition will help to conceptualize independence.

4.5 Contingency Tables

The use of contingency tables can be fruitful in helping upper secondary school students conceptualize conditional probability and independence, Batanero, Estepa, et al. 1996. It may seem a little strange, since contingency tables are a representation of data, and therefore, it is not an object that belongs to probability theory. However, this statistical tool can be used as an opportunity for students to put their probability theory knowledge and conceptualizations into play in a scenario with concrete values and in a way that is descriptive of something that has occurred. In this way, there is a distance to the comparison with divination, which students can sometimes find themselves making. This means that the students have the opportunity, using contingency tables, to work with probability theory without a large amount of preconceptions. Even if they work with frequencies and relative frequencies rather than probabilities and conditional probabilities. In contingency tables, it is easy to determine relative frequencies. Relative frequencies are easier to conceptualize, as it is easier to understand how the quantity of a relative frequency is determined, whereas, on the contrary, it is often difficult to see how the quantity of a probability is determined. Interestingly, probabilities can be determined on the basis of relative frequencies, as mentioned in section 4.1. Furthermore, it is interesting that in these situations the students' intuition will often help them rather than mislead them.

Working with relative frequencies can, in this way, help students to conceptualize conditional probability. However, contingency tables can also be used as a tool to increase students' understanding and strengthen their conceptualization of the concept of independence. As mentioned in section 2, the chi-squared test for independence was once part of the mathematics curriculum in Danish upper secondary school. Chi-squared test can be both a test of independence or a hypothesis test. The two types of tests are following the same procedure, but in this thesis the focus will be on the chi-squared test as a test for independence. The test, which is a product of Karl Pearson's work with χ^2 distributions, can be used if there is a contingency table that describes the relationship between two events, which could have more than two possible outcomes. One can prove Pearson's chi-squared test, but since the chi-squared test does not play a major role in the rest of this work, the content of such a proof does not outweigh the complexity of the proof. If someone wants to read a proof of the test (or several), refer to Benhamou and Melot 2018. Instead of proving the test, the test procedure will now be reviewed using an example. The following contingency table describes the distribution of men and women in the Danish parliament from the three largest parties.

	Socialdemokraterne	**Venstre**	**Moderaterne**	
Mænd	33	18	10	61
Kvinder	17	5	6	28
Sum	50	23	16	89

The first thing that must be done, in order to determine whether the gender distribution in the three parties is independent of the party, is to determine expected frequencies. That is what the gender distribution should look like if it were equally distributed in the parties, in relation to how many men and women are elected to the parliament.

	Socialdemokraterne	Venstre	Moderaterne	
Mænd	34.27	15.76	10.97	61
Kvinder	15.73	7.24	5.03	28
Sum	50	23	16	89

Now the following formula is used to determine the test size, Q, where k is the number of cells in the contingency table, f_i is the observed frequency in the i'th cell and x_i is the expected frequency in the i'th cell:

$$Q = \sum_{i=1}^{k} \frac{(f_i - x_i)^2}{x_i}$$

Which means that the test size, Q, from the example can be determined to:

$$\begin{aligned} Q =& \frac{(33-34.27)^2}{34.27} + \frac{(17-15.73)^2}{15.73} + \frac{(18-15.76)^2}{15.76} + \frac{(5-7.24)^2}{7.24} \\ &+ \frac{(10-10.97)^2}{10.97} + \frac{(6-5.03)^2}{5.03} = 1.4282 \end{aligned}$$

Then this is where the chi-squared distribution comes into play. The chi-squared distribution is different according to the number of degrees of freedom, which is the multiple of the number of rows in the contingency table (without the sum row) minus one and the number of columns (also without the sum column) minus one. In the example, the degree of freedom is therefore $(3-1)\cdot(2-1) = 2$.

It is now desired to determine the probability that the random variable that has the chi-squared distribution with $k-1$ degrees of freedom, X, is greater than Q, i.e. $P(X \geq Q)$. This process contains a definite integral with an infinite upper limit (and Q as the lower limit), which means that the students would have to know the concept of limit, which they do not have in Denmark. Therefore, this part of the calculation is left to a CAS tool, in upper secondary school and in this thesis.

$$P(X \geq 1.4282) = 0.4896$$

The meaning of this is that there is a 48.96% chance that the gender distribution in the three parties would turn out like this by random. To make an assessment of this probability, what is called a significance level is usually chosen in advance before doing the test. Normally, one works with a significance level of 1%, 5% or 10%. The calculated probability must be below the significance level before it can be concluded that the parameters are dependent on each other.

4.6 Simulations

In section 4.2 the large number of misconceptions in probability theory were explained, especially within conditional probability and independence. A tool that can be used to help people finding out where they have made fallacies is simulations. A simulation can clarify misconceptions, as it will point towards a different result than the misconception does. Hopefully, it leads to a reconsideration of the conclusion, and may lead to the person with a misconception finding out that the person has made a fallacy, perhaps even that it is due to a misconception.

Simulations can be performed in several different ways. The most basic is by manually repeating the experiment. It could be sitting down rolling a die or flipping a coin 100 times, but it could also be more complicated experiments. An example that is relevant for the rest of this thesis is the Monty Hall problem, which can be formulated as follows:

> Suppose you are on a game show and are given the choice to select one of three doors. Behind one door there is a car and behind each of the other two doors there is a goat. Once you select a door, say No 1 (which is closed), the host, who knows what is behind each door, opens another door (say No. 3), which contains a goat. You are now given the option of changing your selection to door No. 2 or sticking with door No. 1. What would you do? (Batanero, Contreras, et al. 2014, pp. 366)

For this example, manual simulations could be performed, using an equal area spinner with three areas. By spinning the area spinner, a person can select a door completely at random, after which the person can test the strategy of switching doors and the strategy of sticking with the door. It is possible to spice up the experiment by letting a coin toss decide which strategy should be used for each playthrough. If the person does enough simulations of each of the three strategies, the person will be able to see that it is the smartest to change doors every time, and unwise to stick with the door. The coin toss gives an equal distribution between winning or loosing the prize. Such a manual simulation will give people a certainty in the execution of the experiment, providing a strong ballast for the journey towards a correct conception. Compared to the other simulation methods, this method is the only one that provides this increased understanding of the execution of the experiment. According to several researchers, including Shaughnessy 1992, simulations of this form are a rewarding technique for confronting and overcoming misconceptions.

On the other hand, this type of manual simulations is extremely time-consuming. The consequence of this in a time-pressed teaching setting will be that the students would not achieve many repetitions of the experiment without spending an entire lesson on simulation. Furthermore, the simulation process will quickly become boring for the students, as it by its nature will be a monotonous process. Thus, manual simulations will be a time-consuming and patience-demanding investment.

A much faster and less monotonous simulation method is computer simulations. Computers are obviously a useful device for simulations, as they are extremely much faster to perform an experiment than a human being. For classic and well-known problems, such as The Monty Hall problem, there are pre-produced simulation applets on various websites. These applets often work flawlessly and for most people will be plug and play, and thus a shortcut for students in a class to be able to simulate the same amount of repetitions of the experiment as they would be able to achieve in several lessons of manual simulation in a few minutes. By using this type of simulation, students do not gain an increased understanding of the execution of the experiment, but only an idea of the winning chances of the strategies. This means that pre-produced simulations only have the effect of pointing out false conclusions. Some applets have tried to get a little closer to manual simulations by compromising on simulation speed, and presenting small illustrations that perform the experiment in something that gives the illusion of a manual routine. When a person uses these applets, they will see the experiment being performed very quickly one after the other before they get the output numbers, and thus not just a series of output numbers. This is far faster than what a person would be able to simulate manually with an area spinner, but at the same time much slower than a pure computer simulation. Several applets also allow the user to perform the experiment manually by individual repetitions. A problem with using pre-produced simulation applets for teaching is that students do not understand how the simulation works or whether it is trustworthy.

To avoid this problem, students could also produce their own simulation programs. There are several free coding programs that are suitable for coding simulations, e.g. the program R. The code itself is relatively simple if you are used to coding. In Denmark, and many other countries, it is rarely the case that upper secondary school students have experience with coding, and therefore this approach would often be slow and challenging for many students. However, it is worth noting that the process of coding a simulation of a problem such as The Monty Hall problem can be a learning process for students, as it will require a clear understanding of the execution of the experiment and an initial formalization of the problem so that the computer will be able to understand the problem. This could lead to the students being able to make a real formal proof for the solution of the Monty Hall problem. Such a solution could look like the following.

To solve the problem formally, the following three events are defined:

$$\begin{aligned} C:&\quad \text{The player selects the door containing the car.}\\ G:&\quad \text{The player selects a door containing a goat.}\\ W:&\quad \text{The player wins the car.} \end{aligned}$$

Since $C \cap G = \emptyset$ and $W = (W \cap C) \cup (W \cap G)$, C and G is a partition of the sample space, and we can apply the addition rule

$$P(W) = P((W \cap C) \cup (W \cap G)) = P(W \cap C) + P(W \cap G)$$

Using the definition of conditional probability in this form:

$$P(A \cap B) = P(A|B) \cdot P(B)$$

we get

$$P(W) = P(W \cap C) + P(W \cap G) = P(W|C) \cdot P(C) + P(W|G) \cdot P(G)$$

It is known that $P(C) = \frac{1}{3}$ and $P(G) = \frac{2}{3}$. $P(W|C)$ and $P(W|G)$ depends on the choice of strategy. If you choose to stick at the door you chose first, it is clear that $P(W|C) = 1$ and $P(W|G) = 0$, which mean

$$P(W) = P(W|C) \cdot P(C) + P(W|G) \cdot P(G) = 1 \cdot \frac{1}{3} + 0 \cdot \frac{2}{3} = \frac{1}{3}$$

If you choose to switch, it is clear that $P(W|C) = 0$ and $P(W|G) = 1$, which mean

$$P(W) = P(W|C) \cdot P(C) + P(W|G) \cdot P(G) = 0 \cdot \frac{1}{3} + 1 \cdot \frac{2}{3} = \frac{2}{3}$$

In this way, it is formally shown that there is the greatest chance of winning if you switch.

A general problem with using simulations for teaching purposes is that students use the simulation as an argument for a new conclusion, without considering what justifications and which reasoning that are plausible for their new conclusion using simulation. These are important considerations, as it is through these that students can let go of their misconceptions and develop new, more correct conceptions.

5 The Theory of Didatical Situations

The Theory of Didactic Situations is used as a framework for the teaching design that is to illuminate the thesis' research questions. This theory is one of the classic theories within didactics. Therefore, only a short review of the theory will be given in this thesis, while it is encouraged to read Brousseau 1997 if you want the full review of the theory. Guy Brousseau is the main man behind the Theory of Didactic Situations (from now on called TDS).

In TDS it is important to distinguish between personal knowledge and official knowledge. Personal knowledge is the knowledge that lies in each of our minds of thoughts, i.e. the way we understand things, while official knowledge is the way knowledge is represented when knowledge is to be shared or communicated, e.g. in textbooks, scientific articles or when a teacher explains theorems and definitions etc. and writes it on the board. The epistemological basis of TDS is to expand students' personal knowledge and then to formalize it into official knowledge. To do this, TDS proposes that teaching is structured in five different phases, devolution (or instruction), action, formulation, validation and institutionalization. Often there will be overlaps of the phases and the phases do not have to appear in the order in which they are presented. The devolution phase (sometimes called instruction phase) is where the teacher formulates the problem for the students. The teacher use the devolution phase to set up a didactic *milieu*. The teacher have to explain to the students which problem they have to work on, how long the students have to work on it and which materials and devices the students are allowed to use. The devolution phase is always a didactic phase, which means that it is controlled by the teacher. The action phase is the phase where students work with the problem set in the devolution phase. The teacher have to back off and let the students work on their own. Therefore, the teacher can not give any guidance or instructions. This phase will always be adidactic, which means that the teacher has no control over the course of the lesson. In the formulation phase, students have to formulate hypotheses and conclusions. They can be both tentative or more complete, depending on how far the students are in their work process or how the teaching is structured (for example, it may be intended that the students have to make a preliminary hypothesis before the teacher devolves a modified didactic milieu) . The students' hypotheses and conclusions can be imprecise to begin with. If so, the teacher can encourage students to try to be more precise. The formulation phase often is adidactic. The validation phase is a central phase in TDS, as it is in this phase where the students get their hypotheses and conclusions validated, and thus the first step towards formalizing their personal knowledge into official knowledge. The validation can take place in several ways. It could, for example, be; that the students can validate each other through joint validation (usually controlled by the teacher) or through direct validation from the teacher. The validation phase is often didactic. The institutionalization phase is the phase where the teacher reviews the formal part of the problem (theorems, proofs, definitions, laws, etc.). Here, official knowledge is the focal point. Therefore, this phase will typically be after the students have had the opportunity to expand their personal knowledge in some of the other phases, before this knowledge is formalized into official knowledge. The institutionalization phase is central, as it is not normally possible or necessary for the students to be able to acquire the entire desired amount of knowledge through action phases. In a normal teaching situation, it will always be the teacher who is responsible for this review, and thus the phase is didactic.

The didactic contract is an important concept in TDS, as its effects have a great influence on teaching structured using TDS. The didactic contract is a metaphorical concept (there is usually no physical contract) that describes a special mechanism in the interaction between the teacher and the students. Both students and the teacher have an expectation that the other party will comply with this informal contract. The students expect the teacher to set up a didactic milieu that is suitable for the students in relation to the students' academic ability, and which will at the same time make them smarter and more skilled. The students are expecting that there are correct solutions to the tasks the teacher gives to the students. At the same time, the teacher expects the students to participate actively in the lesson and that they will try to the best of their ability to arrive at the correct solution. This means that the students must accept the didactic milieu. What is particularly interesting about the didactic contract is that a fundamental paradox is built into the contract. The paradox is that the teacher always knows the answers and solutions to all the technical problems and tasks she asks. The teacher therefore does not ask questions to the students because she does not know the answer (as is usually the case in most other social contexts), but with a didactic purpose, namely that the students should find the answer by themselves. The fundamental paradox is thus a consequence of the didactic contract, which makes the interaction between the teacher and the students unnatural and artificial at certain points. The essential problem with the fundamental paradox is that it hinders learning, as it in its way obstructs the didactic milieu a bit. Therefore, it is central to TDS that the fundamental paradox disappears (or at least disappears as much as possible). Since the paradox comes from the didactic contract, the didactic contract must disappear or, probably more realistically, step into the background. The less the didactic contract fills in teaching, the more freely the students can work in the didactic milieu.

There are some effects that derive from the fundamental paradox that typically occurs when the teacher tries to convince himself and the students that the didactic contract is being respected. This means that the teacher more or less unconsciously tries to help the students arrive at the correct answer to the teacher's question. The effects thus only manage to elude the students arriving at the correct answer, while the reality is that it is the teacher who in various ways makes the students say/repeat the correct answer. One of the effects is called the *Topaze* effect, named after a famous play by Marcel Pagnol, which you can read more about in Brousseau 1997 on page 25. The Topaze effect is that the teacher, in the desire to allow the students to answer her question correctly, lowers the difficulty of the question so far that the target knowledge is no longer reachable. Often this happens in a process where the teacher first asks a question that is too difficult for the student to answer. After that, the teacher lowers the level of difficulty a little bit, but the question is still beyond the student's level of knowledge. Then the teacher lowers the difficulty even more, but not enough for the student to answer yet and so on. This leads to the student eventually just having to answer a trivial question. Another effect derived from the didactic contract is the *Jourdain* effect, named after the main character in another French piece of stage art, namely *le Bourgeois Gentilhomme* by Molière, which is described in further detail in Brousseau 1997 pages 25-26. The Jourdain effect is about the teacher convincing herself and the students that the students have acquired knowledge, without the students having necessarily done anything other than saying a correct result or doing exactly what the teacher had told the students to do. The classic example is from Brousseau's own book:

> The student asked to perform rather strange manipulations with jars of yoghurt or coloured pictures is told, "You have just discovered a *Klein group*".
> (Brousseau 1997, p. 26)

In this way, the teacher and the students think that the students have learned something about Klein groups, which is hardly the case, since the students have simply followed the teacher's instructions without necessarily thinking further about what they have looked at. The last effect that will be presented here in the thesis is *metacognitive shifts*. This effect is about downscaling the formal, official knowledge so that it becomes more informal. This will most often happen when the students do not get to expand their personal knowledge, as is desired and expected by the teacher in advance. As a result, the students do not have the prerequisites to be able to follow the teacher's prepared formalization of the personal knowledge into official knowledge. The teacher therefore chooses to present a more informal version of the official knowledge that does not meet the target knowledge. When TDS is used as an evaluation tool, it is interesting to look for what are called *didactic variables* and *fundamental situations*. A didactic variable is variations in the didactic milieu or in the devolution and institutionalization phases. It is especially interesting to look at variations that preserve target knowledge. There will always be didactic variables when examining how a teaching design works in practice. It can be both interesting to investigate why the variation occurred and what it entailed. A fundamental situation is an adidactic situation where the student acquires personal knowledge that enables a formalization into official knowledge shared with the other students in the class. It is interesting to find fundamental situations, as these situations are the culmination of a didactic milieu, and thus possibly can be used to restate a similar teaching process, so that a didactic milieu can be set up that enables students to arrive at the same fundamental situation.

6 Design and A Priori Analysis

In this section, the design process will be presented. It does not make any sense to separate the a priori analysis and the design process into two separate sections, since the a priori analysis is one of the most important tools in the design process. Therefore, the a priori analysis will be presented alongside the design in this section. If you wish to see the designed tasks, they can be found in appendices A.2 to A.7 in the original danish version. The lesson plans that were given to the teacher will be included in an English version in this section, and the original Danish version can be found in appendix A.1. The teaching design extends over one week with five lessons of 55 minutes each however, about 15-20 minutes should be spent introducing the students to the project and giving the students an initial data collection test at the start of the first lesson of the week, and also about 10-15 minutes should be spent giving the students a final data collection test at the end of the last lesson of the week. These lessons are scheduled so that the first two lessons are on Monday and the last three lessons are on Thursday. The lessons are roughly one after the other. Therefore, I have chosen to structure the teaching so that there are a total of six exercises, where three of the exercises have to be completed on Monday and the last three exercises have to be completed on Thursday. These exercises will be presented below one after the other, and are structured in such a way that there are some general information, followed by the actual plan for carrying out the exercise in a TDS-schedule inspired by the MERIA project Jessen and Winsløw 2018 and finally an a priori analysis and argumentation for the design. But before we start on that, here is the general information about the entire course.

6.0.1 General Information

Target knowledge	A strengthening of students' conceptual and intuitive use of independence in probability theory, as well as making students familiar with the formal definition of independence.
Broader goals	Students have to be able to use definitions for both conditional probability and independence. They have to be able to analyze a probability problem that deals with conditional probabilities. They have to know how to read and analyze a contingency table. They have to be able to avoid making the time axis fallacy. They have to know the difference between pairwise independence and mutual independence.
Necessary mathematical prerequisites	There are no specific necessary prerequisites, but knowledge of general probability theory will be a great advantage, but the course can be carried out without this student competence.

Level	The highest mathematical level in the Danish upper secondary school, which is called A-level. The amount of mathematical maturity can be decisive, therefore it is recommended that students have passed their first year of upper secondary school before the course is taught. This means that the students are 17-19 years old.
Time consumption	5 lessons of 55 minutes. Incl. the time spent on introduction and the comprehension tests, which corresponds to 25-40 minutes in total.

6.1 First Exercise - A Contingency Table

Target knowledge	How a contingency table is structured. A initial conceptualization of the meaning of conditional probability. The difference between $P(A \cap B)$ and $P(A\|B)$.
Necessary mathematical prerequisites	There are no specific necessary prerequisites, but knowledge of general probability theory will be a great advantage, but the course can be carried out without this student competence.
Time consumption	30-40 minutes. (Together with the introduction and the initial data collection test, it will last around the entire first lesson, i.e. 55 minutes.)
Materials available	All students receive a printed version of the problem. Other than that just a pencil and a calculator.

Problem: The table shows fabricated data from a random sample, which have to help clarify the effect of mass testing a population for a certain fabricated disease (could be called ”Divoc”). The columns ”Ill” and ”Healthy” refer to whether the persons were infected with Divoc or not at the time of the testing, respectively. While the rows ”Tested positive” and ”Tested negative” refer to whether the persons tested positive or negative, respectively.

	Ill	**Healthy**	**Sum**
Positive test	22	248	270
Negative test	3	2,227	2,230
Sum	25	2,475	2,500

a What is the probability that a random person from this random sample has a positive test?

b What is the probability that a random person from this random sample both has a positive test and is infected with Divoc?

c What is the probability that a random person from this random sample has a positive test if the person is infected with Divoc?

d What is the probability that a random person from this random sample has a positive test if the person is not infected with Divoc? Compare this with question c, what does it mean?

e What is the probability that a random person from this random sample is infected with Divoc if the person tests positive? What is the probability that the person is healthy if the person's test is positive? Compare this. What does it mean?

6.1.1 TDS-Schedule

Phase	**The teacher's actions incl. instructions**	**The students' actions incl. reactions**
Devolution (Didactical) 3-5 minutes	Present the contingency table for the students and ensure that the students understand the contents of the table. Show the work questions to the students. To find the contingency table and task questions see above, where it is reviewed. Tells the students to work with their side mate. Each pair should have a piece of paper with the contingency table and work questions, which can be found in the original Danish version in appendix A.1.1.	Listens and asks follow-up questions. Receives the paper and prepares for the mateship.
Action and formulation (Adidactical) 10-15 minutes	Observes students without interacting with them. However, it is alright to interact with students who cannot get started or have stopped completely and cannot get started again.	Working on the work questions with their side mate. Write their solutions down their papers.
Formulation and Validation (Adidactical) 7-10 minutes	Through the teacher's observations in the action and formulation phase, the teacher has gained an overview of which solutions the different pairs have. Therefore, the teacher can select different pairs of partners to explain their solutions.	Listens when the teacher or classmates speak. Explain their solutions when asked by the teacher. Actively participates in the plenary session.

Institutionalization (Didictical) 10 minutes	Reviews conditional probability theoretically based on the tasks just completed. Uses Venn diagrams to illustrate probabilities and the restriction of the sample space when conditioning.	Listens and takes notes.

6.1.2 Analysis

In the first lesson, part of the time is spent introducing the students to the project and conducting an initial empirical test of the students' conceptualization of independence and thereby also conditional probability. In the exercise it has been a focal point that it is fairly straightforward. For this reason, the exercise is structured with sub-questions of increasing difficulty. The first two sub-questions do not deal with conditional probability or independence, but are questions that should help the students to be able to understand and orientate themselves in the universe of the exercise. The rest of the sub-questions deal with conditional probability, with a focus on how the conditional leads to a narrowing of the sample space. In this way, the sub-questions act like a guide to conditional probability.

The actual exercise the students have to work on deals with a contingency table. In section 4.5, it is explained how research has described how contingency tables can be useful for conceptualizing conditional probability, and more specifically how conditional probabilities are a narrowing of the sample space. The theme of the exercise is testing for a disease called "Divoc". With the Covid-19 epidemic still in the immediate memory, the whole concept of disease testing will be familiar to the students, so they understand the concepts of testing positive or negative. The students have to work on the exercise that they are given on a piece of paper. The students are supposed to work in pairs. Space has been made on the paper for the students' answers. If you want to see the exact problem formulation, you will find the problem formulation in Danish in appendix A.2

The students are first asked what the probability is that a random person from the sample, which forms the basis of this contingency table, has a positive test. This question does not deal with conditional probability, but aims to give the students an overview of the contingency table and the meaning of the numbers in the table. The students' task is quite simple. They just have to find the two correct numbers and divide on by the other to get a frequency. A student who has understood the structure of the table will easily be able to say that the probability have to be the number of positive tests that they find in the sum row on the right side of the table, divided by the total number of people. However, there could be a few students who choose all sorts of numbers from the table with all sorts of arguments for why they chose those particular numbers. Students with a very poor understanding of probability theory may simply be able to find the number 270 as the number of positive tests and think that this number is the probability they search for. Others will find it difficult to understand the contingency table, and thus understand the meaning of the individual numbers. These students could come up with answers to all conceivable combinations of numbers from the table, e.g. $\frac{22}{25}$, $\frac{22}{270}$, $\frac{248}{2{,}475}$, $\frac{22}{25}+\frac{248}{2{,}475}$, ect. Students who lack understanding of both probability theory and contingency tables could find themselves answering 22, as it is the number closest to the cell where positive tests are written. However, the expectation is that almost every student are able to determine the correct probability (which actually is a frequency) or at least arrive at this correct conclusion in a discussion with their partner.

The next question for the students is to determine the probability that a random person from the sample is both sick and tested positive. As in the first question, this does not deal with conditional probability, but have to create an overview of the contingency table. The question is somewhat informal way of asking students to determine $P(\text{Ill} \cap \text{Positive test})$, which is not a difficult task if the student understands the structure of the contingency table, unless there should be some kind of conjunction fallacy described in section 4.2.4. As before, students with a lack of understanding of probability theory and/or contingency tables can come up with all sorts of solutions, which can be a number from the table or a number from the table divided by another number from the table. In this question, some students can incorrectly narrow down the sample space and divide by either 25 or 270 instead of dividing by the entire sample space, 2,500, as they consciously or unconsciously work with it as it was a conditional probability.

In the third question, students' work with conditional probability starts. They have to determine the probability that a random person from the sample will test positive if the person is actually ill, which is called the sensitivity. This time the students have to narrow down the sample space, which of course again requires an understanding of the contingency table, but also an understanding of the linguistic design of the question, which, according to section 4.4, is often really difficult. For students who are not used to working with conditional probabilities, it takes a lot of concentration to understand the nature of the question. Therefore, it is likely that some of the students will misunderstand the question and determine a different conditional probability, most likely the opposite probability (fallacy of the transposed conditional, see section 4.2.3), namely the probability that a random person from the sample is ill if the person has tested positive (the positive predictive value). The students' results are relative frequencies more than they are conditional probabilities, but as was described in section 4.5 it still gives students a basic conceptualization of conditional probabilities.

The fourth question should give the students the opportunity to assess the quality of the test. Therefore, they have to determine the probability that a random person from the sample has a positive test if the person is healthy. The calculation itself is parallel to the one in the previous question, so students may have the same problems. The interesting thing about this task is that the students are asked to compare their result from this task with the result from the previous task. Those students with the correct results will see that the probability of testing positive is much higher for those persons who are ill than those who are healthy. The students are then asked what this means. It is clear to the students that these results imply that the test works and it seems to be good. Students with incorrect results will be able to draw other conclusions, and from this they may see that they have done something that does not hold, and thus reconsider their solution..

In the last question, students have to first determine the probability that a random person from the sample is ill if the person has a positive test, which is called the positive predictive value. Then they have to determine the probability that a random person from the sample is healthy if the person has a positive test and compare these results. Again, the calculations are parallel to the previous conditional probability calculations, but this time the students will get a result that may surprise them. Students will see that the probability that a person from the sample is ill if the person has a positive test is relatively low, namely $\frac{22}{270} \approx 8.15\%$, while the probability that a person from the sample is healthy if the person has a positive test is relatively high, namely $\frac{248}{270} \approx 91.85\%$. This fact can be difficult for some students to understand as they think it seems strange. They believe that this contradicts their conclusion from the previous question, where they concluded that the test is good. Some students will think the test is no longer good, since you are most likely healthy even if you have a positive test. Other students will consider what makes a good test. Obviously, a test that almost always turn out positive when a person is sick and almost never turns out positive when a person is healthy is the most favorable test, but what has most importance; that the test turns out positive for all person with the disease or that it does not turn out positive for all the healthy ones? Some students may conclude that the high probability that a person with a positive test is healthy follows from the fact that it is generally much more likely to be healthy than to be ill, and thus avoid the base rate fallacy. As in the other questions, students may come to different conclusions if they have not conceptualized the structure of the contingency table. The students' work will thus be able to open up a discussion about the quality of the test, which could further open up an initial discussion about independence.

Students who have difficulty with the concept of contingency tables will have continuous challenges in solving all the tasks in this exercise. Therefore, the initial devolution of the teacher is extremely important in order not to leave a small group of students without a chance to work with the exercise. The teacher have to explain to the students how a contingency table is structured and how to understand the meaning of the individual numbers. Furthermore, students work in pairs, which means that a student who has not understood the structure of the contingency table has a partner who can help provide an understanding of this structure. At the end of the exercise, the teacher have to institutionalize the most important points of the exercise. Here, the teacher have to focus on the concept of narrowing down the sample space when working with conditional probabilities. This have to be done visually and illustratively with the help of Venn diagrams, since the students only know a little about the notation of set theory.

6.2 Second Exercise - Another Contingency Table

Target knowledge	A strengthened conceptualization of the meaning of conditional probability. A beginning conceptualization of the meaning of independence.
Necessary mathematical prerequisites	There are no specific necessary prerequisites, but knowledge of general probability theory will be a great advantage, but the course can be carried out without this student competence.
Time consumption	25-30 minutes.
Materials available	All students receive a printed version of the problem. Other than that, just a pencil and a calculator.

Problem: The table shows fabricated data from a random sample, which have to help clarify the effect of mass testing a population for a certain fabricated disease (could be called ”Divoc”). The columns ”Ill” and ”Healthy” refer to whether the persons were infected with Divoc or not at the time of the testing, respectively. While the rows ”Tested positive” and ”Tested negative” refer to whether the persons tested positive or negative, respectively.

	Ill	**Healthy**	**Sum**
Positive test	15	922	937
Negative test	25	1,538	1,563
Sum	40	2,460	2,500

Is this test good or bad and why?

6.2.1 TDS-Schedule

Phase	**The teacher's actions incl. instructions**	**The students' actions incl. reactions**
Devolution (Didactical) 2-3 minutes	Presents the contingency table, which is seen above this schedule or in appendix A.1.2, where it is in its original form. Explains that the contingency table describes another test for the disease "Divoc" that they worked on in the last exercise. Gives students the associated working question; "Is this test for Divoc a good test? Why/why not?". Divide the students into groups of 4-5.	Listen and then join their groups.
Action (Adidactical) 7-9 minutes	Observing the students. Don't interact with them unless it is necessary to get a group working.	Working on the work question.
Formulation (Adidactical) 3-4 minutes	Divide the board so that there is an area for each group and tell the students that each group should write their solution on the board. After this, the teacher continues the observation of the students.	Students write their solution on the board.

Formulation and Validation (Adidactical) 6-7 minutes	Select a group that can be the first to present their proposed solution. The group is chosen on the basis of the observations the teacher has made in the previous phases. This solution can then be discussed in plenary or a new group can be chosen who can present their solution, if this is considered better for the students.	Some students present to the others who listen attentively. Actively participates in plenary discussions.
Institutionalization (Didictical) 5-6 minutes	Reviews independence theoretically based on the problem. Explains definitions for independence, both the definition with and without the use of conditional probability.	Listen and ask follow-up question if necessary.

6.2.2 Analysis

The second exercise is very similar to the first exercise and can therefore be completed significantly faster. Therefore, it is expected to last only about half of the second lesson. If you want to see the original Danish exercise design you can find it in appendix A.3. The difference between this exercise and the previous one is that this time the students themselves have to assess whether a test is good or bad based on a contingency table similar to the table in the first exercise, but with new numbers. If the chi-square test was still part of the curriculum (which it is not, see section 2) it would be a smart method of assessing the quality of the disease test. The students do not get the guiding sub-questions, as in the previous exercise, to guide them towards a conclusion. Another difference between the design of this exercise and the previous exercise is that this time the students have to work on the task in groups instead of in pairs. This is to increase the chance that more students will have interesting and concept-strengthening discussions, since in any group there will be more opinions to the exercise than if the students had to work in pairs.

The expectation is that the students will determine the relative frequencies of testing positive given one is ill and testing positive given one is healthy. If they do this they will notice that the probabilities are close to each other,

$$P(\text{Positive}|\text{Ill}) = \frac{15}{40} \approx 37.5\% \quad \text{and} \quad P(\text{Positive}|\text{Healthy}) = \frac{922}{2,460} \approx 37.5\%.$$

From this, students will be able to conclude that the test is not good, as there is approximately the same percentage that tests positive regardless of whether you test a sick or healthy person. The hope of this exercise is to get the students to use expressions and concepts related to independence, as a linguistic argument. It could, for example, be that they said "whether you are ill or healthy does not affect the outcome of the test", "whether you are ill or healthy has no influence on the outcome of the test" or something similar. There may also be students who focus on the opposite conditional probabilities, i.e. $P(\text{Ill}|\text{Positive})$ and $P(\text{Healthy}|\text{Positive})$, which in this case are

$$P(\text{Ill}|\text{Positive}) = \frac{15}{937} \approx 1.6\% \quad \text{and} \quad P(\text{Healthy}|\text{Positive}) = \frac{922}{937} \approx 98.4\%.$$

If the students try to conclude the quality of the test based on these results, their conclusion will probably be that the test is bad, since there is a very high probability that you are healthy even if you have a positive test. However, in the validation phase of the previous exercise, the students discussed how this high percentage of false positives is due to the fact that the percentage of person with the disease of the entire sample space is very low, which may help students avoid the base rate fallacy. For this reason, some may have stopped themselves from drawing this conclusion based on this result. Students who still have not understood the structure of contingency tables will have as much difficulty with this exercise as the previous one. However, there should not be many of these students left, as they have already had a whole exercise to familiarize themselves with it, and that they now also have a group to support them in their conceptualisation.

In the institutionalization phase of this exercise, it is important that the teacher links the students' work to independence. Students do not know much about independence beforehand. Therefore, the teacher have to define independence, in the way that definition 4.2 does it, which is possible since he defined conditional probability in the institutionalization in the first exercise. However, the teacher should still repeat the definition of conditional probability before the definition of independence so that it is fresh in the students' memory.

6.3 Third Exercise - The Urn Problem

Target knowledge	A strengthened conceptualization of the meaning of conditional probability. A strengthened conceptualization of the meaning of independence. Knowledge of the time axis fallacy, and thus the skills to avoid it.
Necessary mathematical prerequisites	There are no specific necessary prerequisites, but knowledge of general probability theory will be a great advantage, but the course can be carried out without this student competence.
Time consumption	25-35 minutes.
Materials available	Just a pencil and a piece of paper.

Problem: There are four marbles in an urn. Two of them are white and two of them are black.

a You draw a marble from the urn and see that the marble is white, you put the marble in your pocket. What is the probability that the next marble you draw is white, too?

b This time you draw a marble, but do not look at the marble before pocketing it. You draw a new marble, which you see is white. What is the probability that the marble in your pocket is white, too?

c Are the events first marble is white and second marble is white dependent or independent?

Additional task: You have the same urn with the same four marbles in it. This time you draw a marble and you see it is white. Then you put the marble back in the urn. Now you draw again, what is the probability that this marble is white? Repeat the experiment where you draw a marble, which you show to your friend without you seeing it, after that you just put the marble back in the urn. Now you draw a new marble that shows to be white. What is the probability that your friend saw a white marble? Are the events first marble is white and second marble is white this time dependent or independent?

6.3.1 TDS-Schedule

Phase	**The teacher's actions incl. instructions**	**The students' actions incl. reactions**
Devolution (Didactical) 4-5 minutes	Explains the urn problem to the students. The experiment is carried out in front of the students while the task is explained. Explain to the students that they have to continue in their groups.	Listen and ask follow-up questions if necessary.
Action (Adidactical) 8-10 minutes	Observes students and interacts only if it is necessary to get students to start working.	Work on the problem in their groups.
Formulation (Adidactical) 3-4 minutes	Tells students to write their answers on the board. Then the students are observed again.	Write solutions on the board.
Formulation and Validation (Adidactical) 8-10 minutes	Select a group to present their solutions. Then another group can be chosen to present their solution if they have a different explanation or a discussion can take place in plenary. This depends on the teacher's observations.	Present their solution if they are selected by the teacher, otherwise listen. Actively participates in plenary discussions.
Institutionalization (Didictical) 3-4 minutes	Explain to students how conditional probability and causality are not the same, and explains how conditional probability can be used opposite the passage of time. If there is time, the teacher can present Bayes' theorem.	Listen and ask follow-up questions if necessary.

6.3.2 Analysis

The third exercise, called ”the urn problem”, is a well-known problem which are included in both Batanero and Sanchez 2005 and Falk 1989, and more, where they formulate the problem as follows

> An urn contains two white balls and two red balls. We pick up two balls at random, one after the other without replacement. (a) What is the probability that the second ball is red, given that the first ball is also red? (b) What is the probability that the first ball is red, given that the second ball is also red? (Batanero and Sanchez 2005, pp. 251)

In this thesis, it has been chosen to reformulate the problem, for the sake of the students' understanding of the problem. These are not major changes that do not change the basic idea of the problem, but a question has been added to the problem about whether the events are independent, and furthermore a few additional tasks that are relatively simple. In these additional tasks, the students have to do exactly the same thing, but the experiment is carried out with replacement. If you wish to read the problem in Danish as it was given to the students, it can be found in Appendix A.4. To ensure that all students understand the problem, it is presented to the students by the teacher using an urn (or something similar, such as a box) and four marbles (two white and two black, or whatever can be found). With the help of these remedies, the teacher can perform the experiment with its variations a few times, and thus performs a simulation of the experiment, which, as described in section 4.6, helps the students to understand the execution of the experiment and thus make it easier for them to analyze the problem. The students are then ready to work on the problem. The first part will not cause many students trouble, since the task will seem very straightforward to the students whose intuition will be correct in the first place, since they are used to work with problems parallel to this one. The interesting part of the problem is part (b), where the students' intuition may not be correct in the first place, due to the time-axis fallacy. Presumably, on the basis of this intuition, the majority of students will think that the probability have to be $\frac{1}{2}$, since there were two white and two black marbles in the urn when the first marble was drawn. This assumption was tested in Falk 1989 and Falk showed that this fallacy, called the fallacy of the time axis, occurs frequently among students at this level, and is described in further detail in section 4.2.1. It may be that there are a few students who will be able to see this problem in the right context, and thus create a solution equivalent to the solution to the first part of the problem. These students will probably argue for their solution by saying something like if a white marble is drawn in the second draw, then we know that the particular white marble was not drawn in the first draw, therefore the possibilities are only one white marble and two black marbles, which means the probability have to be $\frac{1}{3}$ (which is a correct linguistic argument). After this, the students have to deal with whether the events are independent. Students who have the previous part of the exercise correct will not have much difficulty in concluding that the events are dependent, since their correct analysis of the problem makes the dependence quite explicit. Students who do not have the previous part correct will probably be very uncertain, as they will clearly think that the first draw will be independent of what is drawn in the second draw, but can also see that the second draw is dependent of what was drawn in the first draw. Some students may want to reconsider their probabilities. It is certainly a reason for great uncertainty about the concept of independence, and therefore it is important that the teacher rehearses the definition of independence with the students in the subsequent institutionalization phase and makes the students develop their personal knowledge through the validation phase before that.

Students have to work in groups in this exercise, which brings with it the possibility that there will be groups where students have divergent intuitions. When students with different intuitions have to collaborate, it can lead to interesting discussions between the students, as they will try to convince each other that their intuition is the correct one. The reason these discussions are so interesting is that the students participating in the discussion are highly motivated to provide the strongest and most accurate arguments for their particular intuition. This makes it possible to hear and see students' arguments in a relatively clean form, which is quite worthy for a thesis like this. In addition, this argumentation process will also be very rewarding for the students, as they will experience a more realistic way of working with mathematics, and probably the correct intuition will end up standing as the true one for the students after the discussion. At the end of the exercise, the teacher have to institutionalize the difference between causality and condition. In addition, the teacher also have to institutionalize both Bayes' theorem and repeat the definition of independence. It makes sense to institutionalize Bayes' theorem at this point, since students have just worked on their intuitive understanding of conditional probabilities and this will allow students to subsequently determine conditional probabilities.

6.4 Fourth Exercise - The Monty Hall Problem

Target knowledge	A strengthened conceptualization of the meaning of conditional probability. The power of simulations.
Necessary mathematical prerequisites	There are no specific necessary prerequisites, but knowledge of general probability theory will be a great advantage, but the course can be carried out without this student competence.
Time consumption	55-75 minutes.
Materials available	Laptop, a pencil and a piece of paper.
Problem: Suppose you are on a game show and are given the choice to select one of three doors. Behind one door there is a car and behind each of the other two doors there is a goat. Once you select a door, say No 1 (which is closed), the host, who knows what is behind each door, opens another door (say No. 3), which contains a goat. You are now given the option of changing your selection to door No. 2 or sticking with door No. 1. What would you do? (Batanero, Contreras, et al. 2014, pp. 366)	

6.4.1 TDS-Schedule

Phase	The teacher's actions incl. instructions	The students' actions incl. reactions
Devolution (Didactical) 8-10 minutes	Gives students an ultra-brief sketch of the story behind the Monty Hall problem. Explain the rules of the game to the students. Play through a few examples where students can help develop strategies (e.g. always start with a specific door, switch doors, or stay at the door). If the students themselves do not mention strategies about the choice of the first door, this part is omitted for efficiency.	Listen, participate in the plenary session and pose any follow-up questions.

	If they mention these strategies, it is necessary to exclude them. Tells the students to decide whether they think a strategy would be beneficial and if they think so; which one. Share link to website. Explains to the students that they will now have 5 minutes to play the Monty Hall game on their own. (There is a part of the shared website where you can be the "quiz participant".) Tells the students that they have to choose for themselves whether they play with a strategy or on pure intuition. If they play with a strategy, they must of course also choose which strategy they use. The website does its own statistics.	
Action (Adidactical) 6-7 minutes	Observes the students. Does not interact with them unless it is necessary to get a student started on the work.	Plays the game independently.
Formulation (Adidactical) 6-7 minutes	Divides the students into groups of 4-5 students. Asks students to compare their data, then formulate a preliminary conclusion about how to play the game to have the best chance of winning, and write it on the board. Then the teacher observes the students.	Compare their data with each other, to get the best background to formulate a preliminary conclusion on how to play the game to have the best chance of winning. Write their preliminary conclusion on the board.

Formulation and Validation (Adidactical) 8-10 minutes	Gets a few of the groups to present their preliminary conclusion and why they have come to that conclusion.	Present conclusions and arguments or listen if they do not present.
Devolution (Didactical) 2 minutes	Explains to the students that on the website they can also simulate many repetitions of the game in a short time. Asks the students, just like before, to first work independently with the simulation and then join their groups and discuss their (perhaps) new conclusion, write it on the board, and consider how it could be like this. They must also write their arguments on the board.	Listen and ask follow-up questions if necessary.
Action (Adidactical) 3-5 minutes	Observes the students. Does not interact with them unless it is necessary to get a student started on the work.	Works with the simulation independently.
Formulation (Adidactical) 10-12 minutes	Observing the students. Do not interact with them unless necessary to get a group working.	Working in groups to formulate their conclusion on which strategy gives the greatest chance of winning, write it on the board below their previous conclusion (it doesn't matter if it's the same conclusion).

		Discuss internally in the group how it might be that their conclusion is true. Write their arguments on the board below their conclusions.
Formulation and Validation (Adidactical) 7-10 minutes	Have a group or two explain their conclusion and their arguments to the rest of the class. Opens up the discussion in plenary for other arguments.	Explain conclusions and arguments, listen and participate in the class discussion
Institutionalization (Didictical) 7-10 minutes	Reviews the formal solution and possibly a less formal solution where counting trees are used.	Listen and ask follow-up questions if necessary.

6.4.2 Analysis

In this problem, students have to work with a classic in paradoxical probability theory problems, the Monty Hall problem, and they have to use a precoded simulation applet, whose advantages and disadvantages are described in section 4.6. The problem has been described countless times in previous research, e.g. Batanero, Contreras, et al. 2014. To help students understand the problem, the teacher have to explain in the devolution phase to the students how to play the game. This is done by using the website MathWarehouse 2014, that allowed the user to play the Monty Hall game manually. Here, the teacher have to get the students to come up with strategies which can be used when playing this game. This is done in a plenary session, where the students come up with different strategies. It is expected that the students can come up with the two important strategies to be able to solve this problem, namely ”to change the door” or ”to stay at the door”. However, it is also a possibility that students will mention that it could be a strategy to choose a specific door in the first choice, e.g. door nr. 1. This could be tested by the students, since there could also be learning in this, but since the students should use the simulation applet from the website, MathWarehouse 2014, and since this simulation function is made so that it counts the number of wins and losses for each of the strategies "to change the door" and "to stay at the door", it can be confusing for students if they have to test a strategy that is not part of the program's data collection function. Therefore, already in the devolution, the teacher have to discuss with the students whether it makes a difference to the chances of winning, which door is chosen at the beginning. Here the teacher can help the students on their way by explaining to them that a game show host (e.g. Monty Hall (who funnily enough never included this game in his game show)) does not have a preference for how cars and goats are distributed, this means that there is exactly the same probability that the car is behind one of the doors as it is behind one of the others.

Students should start by playing the game for a few minutes. MathWarehouse 2014 has a feature where you can manually play the game by clicking on doors, that feature the teacher used in the devolution phase. This is so that the students first and foremost try out the game so that they understand the rules of the game. In addition, the students also have the opportunity to test their intuition about which strategy is the best (if any). They will not be able to get a huge amount of data, which means that they will be able to get data pointing in all possible directions. After students have looked at the Monty Hall game individually, they have to discuss in groups which strategy is best (if any). Because the students have worked on the problem individually first, all the students have an opinion on which strategy is best, before they are influenced by each other's views. It can lead to several fruitful discussions, especially since students may have very different conclusions. It is possible that there are students who know about the problem in advance, as it is a very well-known problem that e.g. can pop up on various media the students use (YouTube, Facebook, Instagram, etc.). This does not do much for these students' learning, as they have hardly gone into the depth of the problem, and probably rather have a slightly superficial way of thinking, which, however, means that these students know the right answer. To prevent the students who know the answer in advance from revealing the final point of the problem, the teacher have to ask these students not to reveal anything to the other students and just play along.

In the first group work, the students have to discuss the problem. There will probably be students with different ideas about which strategy gives the best chance of winning. Some (and properly many) students will think that it does not matter if you change doors or not, since when the choice has to be made, there are only two possible doors to choose from, one with a car behind it and one with a goat behind it, therefore there have to be equal probability to win whether you switch doors or not. It is also possible that there are students who can see that the probability have to be conditioned by the first choice, especially because they will pay extra attention to conditional probabilities, as they know it is one of the main subjects within the course. They will hardly be able to give a formal explanation, but say something like that the probability of hitting a goat in the first choice have to be the greatest (2/3), this means that it have to be smartest to change doors when you get the possibility of it. Instead of using intuition, a student could also think of looking at the data the participant has just produced. This will probably come quite naturally to many of the students. The students' conclusions here will of course depend on their data, and their argument for strategy will therefore be "because my data says that is how it is".

The students have to write their preliminary conclusion on the whiteboard and then there is a validation phase, where some of the students are allowed to explain the arguments for their conclusion to the rest of the class. Here it is important that the teacher does not lead the students in the direction of the correct solution by not giving comments on whether groups are right or wrong. This is important as the students have to continue to work on the problem. Now the students have to use the simulation function of MathWarehouse 2014, where they can play through 1000 games in almost no time with one of the strategies. As before, students have to first work with the simulations independently so that they have time to think about their data before returning to their group and discussing whether their now larger amount of data can help them better to conclude something, and if so, what the arguments for their conclusion could be. After such a simulation, the students will end up with a large amount of data that will consistently and unambiguously point to an advantage of changing the door for everyone. As a result, there will probably be quite a few students who thought it did not matter whether you changed doors or not, who are now forced to reconsider this early conclusion. There may be some of the students who think that they have just been "unlucky" with their data and that it have to be a matter of statistical uncertainty. But when the students come back to their groups, they will see that everyone got the same result from the simulation. Here it will dawn on these students that there can be no question of "accident", but that their early conclusion appears to be incorrect. All the students who had an incorrect conclusion after the previous phases (which is expected to be the majority) have to now try to convince themselves and their group why they were wrong and how it can be that it is an advantage to change doors. This is not an easy task, as they have to put their original intuition on the shelf in order to think freely, if they believe that simumations always is right. At this point it gets really interesting as students know the correct answer but not the reasoning behind it. Therefore, students have to think creatively and try to see the problem in a new light without the filter their intuition put on them before. Quite a few students will probably be able to arrive at the informal argument described above by simply going through the game slowly and remembering the structure of the previous exercises on conditional probability. It will be interesting to find out how many that are able to convince themselves of this argument. Since their argument is informal to a greater or lesser extent, there will certainly be some students who find it easier to convince themselves and others in their group, but there will also be students who will not be as convinced. It may be that there are students who throw themselves into making a formal argument, since they have seen the teacher perform such in the institutionalization phases of the previous exercises. However, the expectation is that the students are not yet at a mathematical level where they will be able to carry out such an argument.

After the groups have discussed their arguments, they again have to write their conclusion and arguments on the board. The teacher has not deleted their previous conclusions, so it will become quite clear that most groups have changed their conclusion. Then some of the groups have to present their conclusion and more importantly their arguments. In the institutionalization, the teacher have to provide a formal argument so that the majority of the students hopefully end up being convinced of the correct conclusion. This argument can be seen in section 4.6. The teacher also have to show the students, which is a more extreme edition of the same game. Instead of three doors, there is 100 doors in this example, one with a car behind and 99 with a goat behind. One have to choose a door, after which the game host opens 98 doors with goats behind, such that there is only two closed doors left, the chosen one and another. This example makes the conclusion much more visible.

6.5 Fifth Exercise - The Drawer Problem

Target knowledge	A strengthened conceptualization of the meaning of conditional probability. A strengthened conceptualization of the meaning of independence. Knowledge of the base rate fallacy, and thus the skills to avoid it. Better abilities of analysing problems with conditional probabilities.
Necessary mathematical prerequisites	There are no specific necessary prerequisites, but knowledge of general probability theory will be a great advantage, but the course can be carried out without this student competence.
Time consumption	50-60 minutes.
Materials available	Just a pencil and a piece of paper.

Problem: You have three drawers with exactly two sections in each. In one of the drawers, there is a silver coin in each of the sections. In another of the drawers, there is a gold coin in each of the sections. In the last drawer there is a silver coin in one section and a gold coin in the other.

a You choose a random section in a random drawer. What is the probability that there is a gold coin in the section you have chosen? Now you open the section and see that there is actually a gold coin in the section. What is the probability that there is a gold coin in the second section of the drawer, too?

b In the example from the first part, are the probabilities that there is a gold coin in the section you choose first and that there is a gold coin in the second section of the drawer you chose a section independent?

6.5.1 TDS-Schedule

Phase	**The teacher's actions incl. instructions**	**The students' actions incl. reactions**
Devolution (Didactical) 2-3 minutes	Explains the first part of the drawer problem. Tell the students that they first have to find a solution by themselves for 2 minutes and then discuss their solutions in groups.	Listen and ask follow-up questions if necessary.
Action (Adidactical) 2-3 minutes	Observes and prepares for the division of groups of 2-3 students.	Works independently.
Action and Formulation (Adidactical) 10-12 minutes	Divides the students into groups of 2-3. Informs students that they must be ready to write their solution on the board with any calculations. Tells them they have 10 minutes. Then the teacher observes the students without interacting	Work in their groups to come up with a solution they agree on. They then write their solution on the board with arguments.
Formulation and Validation (Adidactical) 7-9 minutes	Based on the answers on the board and previous observations, 1-3 groups are chosen to present their proposed solutions. After that and/or between the presentations, the teacher can ask questions to the presenting group or to the plenary.	Present, listen and/or answer questions. Ask if they have a probing questions.
Devolution (Didactical) 2 minutes	Gives students the second part of the drawer problem. Tells the students that the process will be the same as before; first 2 min individually and then 10 min in groups.	Listen and ask follow-up questions if necessary.

Action (Adidactical) 2-3 minutes	Observes the students.	Work on the task separately.
Action and Formulation (Adidactical) 10-12 minutes	Observes the students.	Work in their groups to come up with a solution they agree on. Then they write their solution on the board with arguments.
Formulation and Validation (Adidactical) 6-7 minutes	Based on the answers on the board and previous observations, 1-3 groups are chosen to present their proposed solutions. After that and/or between the presentations, the teacher can ask questions to the presenting group or to the plenary.	Present, listen and/or answer questions. Ask if they have a probing questions.
Institutionalization (Didictical) 7-8 minutes	Explains the formal solutions to both parts of the drawer problem.	Listen and ask follow-up questions if necessary.

6.5.2 Analysis

In this exercise, students have to work on a problem that will be called "the drawer problem" in this thesis. The problem is described in Freudenthal 1973, where it is called the drawers problem of Bertrand. The formulation of the problem is basically the same, but a bit has been added to the problem for didactic and understanding reasons in this thesis. The reason students first have to determine the probability of choosing a section with a gold coin is to guide them a little towards a formal solution using the definition of conditional probability and to help them analyse the problem, but also so that students may find it easier to argue for or against independence later. This task should not cause problems for the students, as it is completely general probability calculus, which they have worked on and made a routine in the past. When the students have to work on the main question of the problem, the majority will probably make the same fallacy. They will analyze the problem incorrectly and forget that you are more likely to have chosen a section in the drawer with two gold coins than a section with the gold coin in the drawer with one gold and one silver coin. The students will think of it as equally likely, and thus the probability that there is a gold coin in the second drawer have to be 1/2, that one has either chosen the drawer with two gold coins or the drawer with one of each. In addition to the fact that there could be a student who had enough mathematical overview to be able to give a correct argument for why the solution is 2/3 and not 1/2. Once again, the students have to start by thinking about the problem individually, so that everyone has the opportunity to think through the problem before they have to discuss it in groups. This is to remedy a possible tendency for a talented student to explain the person's thoughts before the rest of the group members have time to think the problem through. This is important, as all students get the opportunity to form their own thoughts about the problem, and not least because the talented student can easily be wrong, while a less talented student could have some fruitful ideas. As always, the validation phase depends on the students' work, but in this exercise it is not certain that any of the groups will have ended up with the correct solution. If this is the case, the validation should be carried out as usual, but it may be necessary for the teacher to ask critical questions about the students' solutions, which may help the students realize that their solutions do not hold water. If some students manage to realize the fallacy, the teacher (possibly in collaboration with the students) can argue for the correct solution. This may develop into a small institutionalization phase, even if it is not indicated in the exercise schedule. If there are groups with the correct solution and argument, the validation phase can be carried out as usual.

After this, students have to work on the added task on independence. Students who are fully convinced of the solution and the arguments for the drawer problem will not have great challenges in seeing that the events are dependent. Students who are not yet fully convinced, on the other hand, will probably have a greater tendency to stick to their intuition from the first part of the problem. These students could have a "feeling" that the events have to be independent. It is not certain that they can explain why they think so, and since at the same time, due to the didactic contract, they will think that the teacher's solution and arguments have to be correct, this may force these students to think carefully about the problem. On the one hand, their intuition tells them one thing, and on the other hand, the students may have doubts about their own intuition, based on the teacher's (or possibly other students') arguments in the first part of the problem. This doubt about their own intuition will perhaps cause some students to view the task in a more formal way, since in previous exercises the students have seen how a formal approach to problems can help them to the correct solution despite an incorrect intuition. The students have all the necessary tools to be able to show that the events are independent, as they know several definitions of independence, e.g. definition 4.2 and theorem 4.4 from section 4.1, which can be used in this exercise. They have just become more or less convinced that the probability that there is a gold coin in the other section in the drawer where you have chosen a section, given that there is a gold coin in the section you have chosen, is 2/3. As one of the very first things in this exercise, the students determined the probability of selecting a gold coin in a random section selection to be 1/2. The students have to convince themselves that the probability that there is a gold coin in the section they choose is equal to the probability that there is a gold coin in the other section in the drawer where they have chosen a section. When this argument is in place for the students, they will be able to see that the events are dependent, since they do not satisfy $P(A|B) = P(A)$. Of course, there is still the possibility that some students are unable to implement this more formal approach, and therefore may end up with a frustrating feeling that their intuition is probably incorrect, but they cannot come up with a better argument. Since the students have to work in groups (after they have worked individually on the task first, with the same reasoning as above), students with this frustrating feeling will quickly begin a group collaboration, where there will be other students who can help them on the right path. The validation phase is carried out in the usual way and is followed by an institutionalization phase, where the teacher have to institutionalize both the solution to the central problem in this exercise, but also the independence part of the problem together with the other definition of independence (theorem 4.4), which is also called the multiplication rule.

6.6 Sixth Exercise - The Three-Event Problem

Target knowledge	A beginning conceptualization of the meaning of mutual independence and pairwise independence. The students have to know the difference between these concepts.
Necessary mathematical prerequisites	There are no specific necessary prerequisites, but knowledge of general probability theory will be a great advantage, but the course can be carried out without this student competence.
Time consumption	30-45 minutes.
Materials available	Just a pencil and a piece of paper.
Problem: You have two fair coins, one 1 krone and one 2 krone. You want to play heads or tails with both coins at the same time, but before that you have to consider the probabilities of the following three events: the 1 krone becomes a head, the 2 krone becomes a head and exactly one of the coins becomes a head. That is, neither more nor less than one. After these initial considerations, you have to decide whether all three events are independent or dependent.	

6.6.1 TDS-Schedule

Phase	**The teacher's actions incl. instructions**	**The students' actions incl. reactions**
Devolution (Didactical) 3-4 minutes	Explains to students about the 3-event problem. Explains that they have to work on this problem just like the previous two times, first seperately and then in groups.	Listen and ask follow-up questions if necessary.
Action (Adidactical) 2-3 minutes	Observes the students.	Work on the question separately.

Action and Formulation (Adidactical) 10-15 minutes	Observes the students.	Work in their groups to come up with a solution they agree on. Then they write their solution on the board with arguments.
Formulation and Validation (Adidactical) 8-10 minutes	Based on the solutions on the board and previous observations, 1-3 groups are chosen to present their proposed solutions. After that and/or between the presentations, the teacher can ask questions to the presenting group or to the plenary.	Listen and ask follow-up questions if necessary.
Institutionalization (Didictical) 7-10 minutes	Reviews the formal explanation of why the 3 events are not independent when they are pairwise independent. Relates to the binomial distribution.	Listen and ask follow-up questions if necessary.

6.6.2 Analysis

This exercise is the last part of the teaching designs and involves the students working with what is called the ”Three-Event Problem” in the thesis. The exercise is heavily inspired by Székely 1986, pp. 12–13. Once again, a little has been added to the problem, so that the students both find it easier to understand the content of the problem and, in addition, are helped to get started in a specific way that has a didactic purpose. The first two considerations will not cause students any problems. They are used to working with coin tosses in probability theory and knowing that the probability of heads is 1/2 regardless of which coin is played with (not to mention if the coin is fair). The third consideration is almost as easy for the students, but there is a possibility that very few of the students do not remember that there is a difference between whether the outcome is (H,T) or (T,H), where H stands for head and T stands for tail, and the first position in the parenthesis indicates the outcome of the 1 krone, while the second position in the parenthesis indicates the outcome of the 2 krone. This means that students will believe the sample space is (H,H), (H,T) and (T,T) and the only successful outcome is (H,T), which entails a probability of 1/3 instead of 1/2. But as I said, almost all students at this level should be able to determine this probability to 1/2. The reason why the students have to start with these considerations is so that they understand the three different events of the problem.

In this exercise, the students will deal with something that is not clearly defined for them yet, namely the independence of several events. Different from the other tasks, the students' intuition in this problem will probably lead some students to the correct solution, while other students probably will end up with an incorrect solution based on their intuition. Some students want to try formal calculations, but these are not very demanding if you have analyzed the problem correctly. However, the problem analysis is more challenging. There will be some students who, at first glance, will think that if you know the outcome of two of the events, you will be able to determine the outcome of the last event. Thus, the events have to be dependent, since the last event depends on the outcome of the other two. Other students may want to try something similar, where they simply compare two events at a time, instead of all three at once. Eg. could they say that the two events; the 1 krone ends up head and 2 krone ends up head, is independent, as they clearly do not depend on each other. They could say the same about the events the 1 krone ends up head and exactly one 1 krone among the coins and about the events the 2 krone ends up head and exactly one 1 krone among the coins. In doing so, these students will think that they have shown that the events are independent, which they have in an informal way. Once again, students first have to work on the problem individually and then work on the problem in groups. The reason for this is the same as the previous times. In the teacher's institutionalization, students will learn about the difference between pairwise independent and mutual independent.

7 Methodology

When creating a teaching design (such as the one just described in section 6) with the aim of testing its effects in relation to some specific learning objectives, and to test the students' knowledge of independence in probability theory, it is important to think about how you can and will collect data. In addition, it is also important to consider how the entire teaching situation can seem as normal as possible for the students, so that they do not change their behavior too much. How these focal points have been taken into account and what thoughts have been behind them are interesting in a scientific context. Therefore, these considerations will be explained in this section.

To keep the students from feeling that they are taking part in an experiment, their usual teacher will teach them, while the author of the thesis acts as an observer. This has the advantage that students know the teacher beforehand and therefore are used to his personality and their teacher-student interactions. In addition, it also frees the thesis author to focus all efforts on observing, as there is no need to think about teaching. One of the disadvantages of doing it this way is that the teaching process depends on the teacher's adaptation of the teaching materials, i.e. how the teacher understands the lesson plans and how the teacher combines it with the teacher's personal style. To ensure that the teacher understands the lesson plan as intended, a meeting is held between the author of this thesis and the teacher a week before the course is scheduled, where they can talk their way through the lesson plans, so that the teacher is as far as possible aligned with the author of this thesis before the teaching has to be carried out. To prepare the students for the course, it starts with a short introduction given by the author of this thesis, where the students are told that they will participate in a study that will be used as the basis for an academic work that will lead to a final Master's thesis. The students are informed that they must work with independence within probability theory and that they are not expected to know anything about this area in advance. They are also prepared for the fact that the teaching is organized a little differently than they are used to, but that it will still be their teacher who teaches them. They are informed that the author of the thesis will observe them and audio record them, but that everything will be performed completely anonymously, so that neither school, class, student names nor the teachers name will appear in the thesis. In addition, the students are informed that the author of the thesis will ask everyone involved for permission before starting an audio recording, and if someone do not want to be audio recorded, they simply have to decline.

A very central part of the data collection is the tests the students are exposed to before and after the course. Both can be seen in appendix A.8 to A.11. The two tests are not identical, but almost, since there is only a difference in the last tasks, where the students have to decide whether two events are independent or not in three different situations. The situations and the probabilities the students get are different, but the task type is the same. In order to ensure that the students answer the tasks of the test as freely as possible, the test answers are not given to the students' teacher, which the students are told, and that the test will therefore have no influence on their grades. To make it possible to compare a student's first test with the student's second test, all students are asked to write a name on the test. They are told that it does not matter whether it is their own name or a made-up name, as long as they remember to write the same name on both the first test and the second test. So if a student wants to hand in a completely anonymous test, they can e.g. write Mickey Mouse or HC Andersen instead of their own name. The purpose of the test is to test whether the students develop their use of language about independence, whether they achieve certain of the learning objectives of the course and whether they become better at using and become more familiar with the area's notation and formulas. The advantages of making such a test is that it enables the collection of data from all of the students in the class at once. One of the disadvantages of using such a test is that students do not write down all their thoughts. The test is formulated so that it should encourage students to write reasons and arguments for their answers, but this does not mean that students will write all the thoughts that precede their arguments or the arguments they choose not to use if they convince themselves that they are incorrect.

Exactly these thoughts of the students are very interesting, as they tell about the students' process and not only about the end of their process. It has been a clear desire to collect data that can help describe this process, too. In order to collect this kind of data, observations will be made of the students when they are working on the problems. These observations will take place through a combination of classic observations, where the author of this thesis will listen to the students' work and note when something interesting happens, and audio recordings of the students. Since it is not possible to observe all the students at once, a group is selected for each problem on which the observations are focused. It will be the same group that is observed and audio recorded at the same time, so that these two data collection methods can complement each other. The audio recordings will both help the observer, as he do not have to write everything down, but can focus on the most interesting and, as it becomes possible, to use quotes from the teaching verbatim. Another important source of data is the students' solutions to the various problems. Therefore, as far as possible, the students' answers are collected when they write them on paper and pictures are taken of the board when the students have to write their solutions on it. This data form will show how students will present their final solutions and arguments. This shows something about what the students at the end are most convinced of. The teacher's opinions on the teaching design can also be interesting. Therefore, in connection with the previously mentioned clarifying meeting between the teacher and the author of this thesis, the week before the teaching is to be tested, an interview will be conducted with the teacher, where the focus will be on the teacher's expectations for the teaching. In the same way, the teacher will also be interviewed after all the teaching has ended, in order to get the teacher's reactions to the teaching, as well as the teacher's views on how it all went.

8 Results and A Posteriori Analysis

This section is organized in such a way that there will be a general review of the entire teaching process at first. Here, the adaptation of the lesson plan to real teaching will be analysed, as well as the participation of the teacher and the students on the premises of the lesson plan. After this, each individual exercise will be analyzed independently one at a time. Here, the analysis will be based on data collected during the students' group work in action, formulation and validation phases, as well as their solutions and arguments written on the board or paper for the various exercises. After that, the overall results from the students' comprehension tests will be presented. Here, the general results and trends of the class will be looked at, and these observations will be supplemented by quotes and results from individual specific students. At last, the teacher's and the students opinion on the implementation of the teaching and the structure of the lesson plans will be analysed.

But before starting on this, the prerequisites for the teaching will be reviewed. The class that took part in the project were in their second out of three years of upper secondary school, which means they were around 17-19 years old. The class has a field of study with Mathematics and Social Studies at the highest upper secondary school level (A-level). In their first year of upper secondary school, the class had taken a course on probability theory, including work with frequencies, which they had been given a recap of before they had to participate in this project. The class has not been taught statistics or binomial distributions yet.

8.1 General observations

If a teacher's teaching varies from the lesson plan, e.g. by helping the students when not intended or not complying with the phases, it becomes difficult to examine what the purpose was from the beginning. The reason for any deviation from the lesson plan might be of interest, and therefore it can be giving to have the teacher explain the reasoning behind the deviations.

It is challenging to estimate how long each phase will take. There can be many small episodes that are difficult to foresee, which can delay the entire plan. Each phase had a relatively lenghty amount of time in the lesson plan, especially the phases where the students had to participate actively (typically the action, formulation and validation phases). As previously mentioned, there was five lessons of 55 minutes available for the course, which means 275 minutes in total. However, some of the time must be spent on introducing the students to the project and on doing the tests at the beginning and at the end of the course, which means that only about 250 minutes is left for the teaching. The teaching process took less time than expected, when it was tested. About 35 minutes less.

If we look at the individual exercises, only two out of the six exercises were finished faster than expected. The exercises carried out Monday (the first three exercises) were well timed as they could be carried out in the two scheduled lessions. In the third exercise (the one with the urn problem), the teacher spent little time on the devolution. In return, the institutionalization phase took a bit longer than expected. In the exercises carried out on Thursday, namely the last three exercises, there were deviations from the time estimates. The fourth exercise (the one with the Monty Hall problem) was about 15 minutes shorter than expected. The main reason for this deviation was primarily that there was a much larger proportion of the students who already knew about the Monty Hall problem beforehand. The students who were familiar with the problem had dealt with it in primary school or knew it from YouTube-videos. In a break between the first and second exercise, a conversation between a larger group of the class's students was overheard. This group discussed independence and conditional probability based on their experiences from the first lesson. In this discussion, a student explained the Monty Hall problem with associated solutions and arguments. This was an unprompted discussion between the students, and not a discussion that could be predicted. However, it is clear that this discussion must have had an impact on how long the students were going to spend on the problem, as they knew solutions and arguments in advance. In addition, several students were quick to find the website's simulation function in the first action phase. As such, the teacher decided that it did not make sense to go through a long action, formulation or validation phase in the second round, as the students would not be able to come up with anything new when they had already done the simulations. Therefore, these phases largely merged into the first round of phases. The fifth exercise (the one with the drawer problem) showed the largest deviation from the estimated time, as it was about 20 minutes shorter than expected. The exercise was divided into two parts, as the students first had to go through an action, formulation and validation phase, where they worked to determine the probabilities of the drawer problem. This part was well estimated. After this, they had to go through another action, formulation and validation phase, where they worked on part (b) of the drawer problem. It was in this part the estimation was wrong. The different time consumption was due to the fact that most students agreed on independence right away, which meant that they barely discussed their arguments. The validation and institutionalization phases merged and lasted a shorter amount of time than expected, as no time had to be spent discussing different solutions. In the last exercise (the one called the three-event problem), the students spent very short time on the action and formulation phases combined. Later (in section 8.7) it will be explained that the students had great problems with this exercise, and these problems led to the students formulating answers that they themselves were not convinced of. Based on this, the teacher decided it made more sense to continue to the next phases. The validation phase was also a bit shorter than expected, which was because the students were nowhere near the intended solutions or arguments. However, the institutionalization phase was twice as long as intended, a full 20 minutes. The teacher assessed (correctly) that the students found the exercise very difficult and therefore chose to spend a little longer going through this exercise.

The teaching method was new to both the teacher and students. The teacher had encountered the Theory of Didactic Situations a few times during his education, but had not used it as a design tool for his teaching himself. Both the teacher and the students were generally very good at complying with the phases of the teaching method and the associated rules. The teacher was good at refraining from helping the students when they were in the adidactic phases and the students were also good at trying to solve the tasks without seeking confirmation from the teacher. That is that the didactic contract was put in the background, except in some special situations that will be explained and analyzed later. The phases were respected by the teacher, but quite generally throughout the course the validation phase and the institutionalization phase merged together. This is by no means a bad thing, but the teacher tended to start the institutionalization early in the validation phase, after which the arguments became much more directed by the teacher rather than the students themselves.

8.2 First Exercise - The First Contingency Table Problem

In this exercise, the students had to work on the following problem: The table shows fabricated data from a random sample, which have to help clarify the effect of mass testing a population for a certain fabricated disease (could be called ”Divoc”). The columns ”Ill” and ”Healthy” refer to whether the persons were infected with Divoc or not at the time of the testing, respectively. While the rows ”Tested positive” and ”Tested negative” refer to whether the persons tested positive or negative, respectively.

	Ill	**Healthy**	**Sum**
Positive test	22	248	270
Negative test	3	2,227	2,230
Sum	25	2,475	2,500

a What is the probability that a random person from this random sample has a positive test?

b What is the probability that a random person from this random sample both has a positive test and is infected with Divoc?

c What is the probability that a random person from this random sample has a positive test if the person is infected with Divoc?

d What is the probability that a random person from this random sample has a positive test if the person is not infected with Divoc? Compare this with question c, what does it mean?

e What is the probability that a random person from this random sample is infected with Divoc if the person tests positive? What is the probability that the person is healthy if the person's test is positive? Compare this. What does it mean?

The observed pair had an interesting process of answering the problem. They started by answering the tasks incorrectly because of imprecise reading of the subquestions. Which means that they have difficulties with the linguistics. According to the a priori analysis (section 6.1.2), it would be the structure of the contingency table more than the linguistics, that would be a challenge for the students. At some point, the students realized that something was wrong, after which they came to a standstill. They asked the teacher for help. The teacher helped the students by asking them to read the questions again. The students read the questions again, this time more accurately, and then the linguistic difficulties were smaller, and they could answer the tasks correctly.

Part (a) was easy to solve for the observed pair, and the subsequent validation showed that all pairs found part (a) to be easy. This was to be expected, as mentioned in the a priori analysis (section 6.1.2), because the students have old knowledge about frequencies from an previous course about probability, as mentioned in the methodology section (section 7). Already in part (b), the observed pair shows that they have difficulty with the linguistic, and with the structure of the contingency table, too. They do not pay enough attention to reading the task question thoroughly. The following is a quote from the pair's discussion about the answer to part (b)

> Student 1: *If you are sick, do you have to look at the column where it says ill?*
> Student 2: *Yes, and tested positive... Well, no, because you can test negative and be ill. That's 25 out of 2500.*
> (Translated by the author.)

They end up determining the frequency of being ill, which does not take positive (or negative) tests into account. This fallacy is a bit familiar with the conjunction fallacy, which is described in section 4.2.4, since the students mix up $P(\text{Ill}\cap\text{Positive})$ and $P(\text{Ill})$. The students do not return to this question, and they do not find out that they have not found the correct answer until the validation phase, where the majority of the other groups have answered correctly, which still is predictable because of the students old knowledge about probability and frequencies. The pair continued with part (c), where they once again misread the question, and they think they have to answer the same thing as they had to answer in part (b). They come to the conclusion that it must be $\frac{22}{2500}$, which actually is $P(\text{Ill} \cap \text{Positive})$, and had been correct in part (b). This fallacy is like the conjunction fallacy (see section 4.2.4), but reverse, since the students determine $P(\text{Ill} \cap \text{Positive})$ instead of $P(\text{Positive}|\text{Ill})$, and not the opposite. If they had not forgotten what they were asked about in part (b), they might have wondered that they understood part (b) and (c) as the same question.

They make exactly the same mistake in part (d), where they answer $\frac{248}{2500}$, which is $P(\text{Positive} \cap \text{Healthy})$ instead of answering $P(\text{Positive}|\text{Healthy})$, which had been correct. This is the same fallacy as above. In part (d) they are asked to compare their answers from part (c) and (d), too. With the mistakes the pair have made, their results from these tasks are $\frac{22}{2500} = 0.88\%$ and $\frac{248}{2500} = 9.92\%$, respectively, in contrast to the correct results, which is $\frac{22}{25} = 88.00\%$ and $\frac{248}{2475} = 10.02\%$, respectively. A consequence of this is, that the pair's comparison is completely different from the intended one. The pair are not surprised by their results, as seen here:

> Student 2: *Okay, there is at least a much greater chance of having a positive test... but actually being healthy than having a positive test and being ill... But what that means. I don't know what to say.*
> Student 1: *It's probably not that accurate.*
> Student 2: *So people are probably a little more relaxed if they test positive.*
> (Translated by the author.)

What the students see here is basically just the base rate, i.e. that there are far more healthy people than sick people. In principle, the students have the knowledge needed to be able to determine the correct frequencies, but they have problems linking the problem's sub-questions to that old knowledge.

In part (e), they misread the task in the same way once again, which means that they think they are being asked the same thing as in part (c). It causes quite a bit of surprise for the pair, who from their normal teaching (in all subjects) are used to the fact, that no two questions are the same. The students suspected that they had misunderstood something in the previous parts, as it would be a breach of the didactic contract if they were given the same task twice in a row by the teacher. The students return to part (c) to check the differences in the wording of the questions. However, they still believe that they are being asked the same question, although the wording is different in the two questions. Therefore, they try to bring up several different method proposals instead, which mostly have the character of guesswork. Eg. the pair consider whether to add any of the numbers together, but they cannot figure out which ones should be added or why. In the end, the pair gives up completely and just sits with a look of hopelessness and sometimes outbursts some thing like ”*It makes no sense*”.

The pair ends up asking if the teacher can help them, which is a violation of the rules of TDS. The teacher explains to them that he must only help them clarify the task questions and not help them with what they have to do, in trying to comply with the rules.

> Student 2: *I just don't think it makes any sense. Or it is so... because it sounds as if (c) or (d) are the same as (e), other than it is worded the other way around.*
> Teacher: *The wording actually means something.*
> Student 2: *Yes,*
> Teacher: *Slowly and thoroughly.*
> (Translated by the author.)

This is an example of indirect negative validation from the teacher. Which also shows that the didactic contract has not completely faded into the background yet, as the teacher indirectly tells the students that their solution is not correct, and that in order to solve the tasks correctly, they will have to read the task formulations again, focusing of the the formulations.

After these comments from the teacher, the pair starts over with part (c). After repeating their initial result and argumentation, student 2 realizes the actual meaning of the question. Then student 2 calculates problem (d) and compares (c) and (d) completely as intended. Student 2 tries to continue with part (e), but once again misreads the task, so she understands it just like part (c), which is the fallacy of the transposed conditional probability. Again she realizes that something is wrong. This time, however, she reads the questions thoroughly again, and finds out that they are worded the other way around.

> Student 2: *There are 22 tests that are positive... and how many... there are 25 infected... No, because the order is reversed... It can't be 25 out of 22.*
> (Translated by the author.)

From her old knowledge, she knows that a probability cannot have a value greater than one, which causes her to reject the conclusion. After a short period of time in deep concentration, student 2 gathers the threads and determines the correct frequency, which means she got rid of some of the linguistic problems.

Student 2 realizes that student 1 does not understand what she is doing. Therefore, student 2 tries to explain it to student 1. The explanation becomes very incoherent, as student 2 tries to reformulate the questions in the way she understands them.

> Student 2: *That's because it is asked the other way around here. Before it asked; if the person is infected... it's because you should almost have reformulated it... or so... I think... because then it makes more sense. If the person is infected with Divoc... okay how many are infected with Divoc? There are 25 of them. What is the chance that that person tests positive out of the 25 who are sick... What is the chance... How many of them test positive. And there were 22 of them. Now it asks the other way around. If the person's test was positive... okay... how many tests were positive of those that were taken, there were 270... Then you have to go over and look in the other one... over there. How many of them were really sick... There were 22 of them.*
> (Translated by the author.)

It is clear that student 2 reformulates the questions in her head in order to better translate them into what she has to do. For student 2, it is more of a linguistic exercise than a calculation exercise. Student 2 has expanded her personal knowledge and wants to help student 1 expand her personal knowledge, too, in the same way, as she did it. It is clear that it is difficult for student 2 to transfer her personal knowledge to official knowledge, such that it becomes shareable. After this, the students were interrupted by the teacher so that they could start the validation phase, which is a shame, since the sharing of knowledge from student 2 til student 1 was interesting.

In the validation phase, nothing particularly interesting happened, as virtually all students agreed on the correct results. Based on the observation of students 1 and 2, it was the linguistic process from the students' first imprecise reading of the questions to them placing more emphasis on reading them accurately that was the interesting part of this exercise.

In the institutionalization phase, the teacher started by reviewing the definition of conditional probability based on the solutions from the problem. The teacher told the students that from their solutions it appeared that a conditional probability could be determined by $P(A|B) = \frac{|A \cap B|}{|B|}$, explaining the notation continuously, as the students did not know the notation for conditional probability in advance. The teacher explained that they had worked with a symmetric probability space, and since the students knew that their frequency approach only worked for symmetric probability spaces, he wanted to see if he could transform the frequencies into probabilities. Therefore, the teacher shortened the fraction by $|S|$, i.e. the size of the entire sample space. Since $P(A \cap B) = \frac{|A \cap B|}{|S|}$ and $P(B) = \frac{|B|}{|S|}$, it must hold that $P(A|B) = \frac{P(A \cap B)}{P(B)}$. The teacher informed the students that this definition could be extended to apply in all probability spaces. After that, the teacher illustrated how a conditional probability is a probability in a narrowed outcome space, using Venn diagrams. The teacher drew a set ball he called S, which should illustrate the entire sample space in a symmetric probability space, then he drew two set balls of different sizes, A and B, illustrating the successful outcomes for two different events. The two set balls overlapped each other. Then the teacher showed how all the results from the problem could be illustrated using these set balls. The teacher focused on how conditional probability was a narrowing of the sample space.

8.3 Second Exercise - The Second Contingency Table

In this exercise, the students had to work on the following problem: The table shows fabricated data from a random sample, which have to help clarify the effect of mass testing a population for a certain fabricated disease (could be called "Divoc"). The columns "Ill" and "Healthy" refer to whether the persons were infected with Divoc or not at the time of the testing, respectively. While the rows "Tested positive" and "Tested negative" refer to whether the persons tested positive or negative, respectively.

	Ill	**Healthy**	**Sum**
Positive test	15	922	937
Negative test	25	1,538	1,563
Sum	40	2,460	2,500

Is this test good or bad and why?

This exercise is somewhat similar to the first exercise, which should enhance the students' abilities of the exercise, though this exercise is more open than the previous one. However, the students approached the task differently than expected. In the validation phase, it could be seen that the groups had focused on a single relative frequency (e.g. the sensitivity or specificity), and not compared several frequencies, as in the first exercise. There were five groups in total, and they all argued that it was a bad test. Two of the groups used the sensitivity, there were only $\frac{15}{40} = 37.5\%$, to argue, that the test was not good. One of the groups pointed out that $\frac{922}{937} = 98.4\%$ of the positive tests were false positives, which is the compliment to the positive predictive value. Another group expanded on this argument when they argued that the test was bad by explaining that there were many false tests (meaning both false positives and negatives). The last group wrote this on the board: "*The test is not good because the number of ill and healthy is very different from the number of positive and negative*" (Translated by the author). This argument is quite close to the previous one. None of the groups followed the same series of calculations as in the first exercise, which was otherwise what was expected in the a priori analysis. If the students had followed parts (c), (d) and (e) from the first exercise, they would have found that the frequency of the event that a person is ill and the frequency that a person tests positive are equal. One of the goals of this exercise was to get the students to compare the two relative frequencies, $P(\text{Positive}|\text{Ill})$ and $P(\text{Positive}|\text{Healthy})$, such that they could see, that the test was equally likely to give a positive result regardless of whether one was sick or healthy.

In the observed group, there were several misunderstandings of the problem. First of all, the group believed that the sample study consisted of the same people as in the sample study of the first exercise, even though the number of ill/healthy was not the same as in that exercise. This indicates a lack of clarity in the formulation of the problem, and caused the group to spent a lot of time discussing (based on covid-19) how the tests in the two exercises could be different. This led to a discussion about how different tests work, and then whether it is possible to be more or less infected with covid-19. Of cause, none of this was intended for the exercise. The observed group was the group that ended up writing about a high probability of false tests on the board. Two of the group members quickly started suggesting ways in which they could better assess the quality of the test.

> Student 3: *I don't know if it makes a difference, but should we try writing it up as a percentage? Just to get a little idea about it... So out of them...*
> Student 4: *We can again look at how many of the positive test that are actually sick. Then we can get an assessment of how good... how much it actually works when it gives a positive test.*
> (Translated by the author.)

Student 3's request is a little unclear in intention. Does he want to write all the numbers in the contingency table into percentages or does he want to calculate conditional probabilities just like in the first exercise? It is clear that student 4 will look at the positive predictive value, which would also be more useful than rewriting all the numbers as percentages of the entire sample study. Before the group determines some of the conditional probabilities, student 4 begins to explain what seems like a relatively immediate assessment of the contingency table.

> Student 4: *There are far more healthy people than sick people. Because then it explains both. This explains both that there are many healthy people with a positive test and that there are many healthy people with a negative test.*
> (Translated by the author.)

Here student 4 avoids making a base rate fallacy. The base rate fallacy is described in section 4.2.2. Student 3 reports that he does not understand what student 4 means. Before student 4 can answer, student 5 breaks in.

> Student 5: *There are also many more who are positive in this sample... There were only 270 who had tested positive before.* (In the first exercise.)
> Student 4: *Yes, more people have tested positive... And they have less ill people.*
> (Translated by the author.)

Here, the group realizes that there are many more positive tests (937) than in the first exercise (270), and that there are fewer people who had a true positive test (15) than in the first exercise (22). The group then comes to the conclusion that there have to be many more false positive tests. The group concludes in exactly the same way that there are also many false negative tests. From this they conclude that the test is bad and their discussion switches to the previously mentioned irrelevant discussion.

The validation phase developed into a discussion about sensitivity, specificity and positive/negative predictive values, although without the technical terms. In the institutionalization phase, the teacher explained that if you looked at e.g. the probability of being ill, given a positive test, $P(\text{Ill}|\text{positive}) = \frac{15}{937} = 1.6\%$, and then on the probability of being ill, $P(Ill) = \frac{40}{2{,}500} = 1.6\%$ then you can see that they are the same. The teacher explained that this meant that the event of being ill and the event of being tested positive are independent. He explained that the definition was actually that $P(A \cap B) = P(A)P(B)$ should hold. He showed students how to derive the second definition, with conditional probabilities, using this definition and the definition of conditional probability, which the students had seen at the end of the first exercise $P(A|B) = \frac{P(A\cap B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A)$.

8.4 Third Exercise - The Urn Problem

In this exercise, the students had to work on the following problem: There are four marbles in an urn. Two of them are white and two of them are black.

a You draw a marble from the urn and see that the marble is white, you put the marble in your pocket. What is the probability that the next marble you draw is white, too?

b This time you draw a marble, but do not look at the marble before pocketing it. You draw a new marble, which you see is white. What is the probability that the marble in your pocket is white, too?

c Are the events first marble is white and second marble is white dependent or independent?

Additional task: You have the same urn with the same four marbles in it. This time you draw a marble and you see it is white. Then you put the marble back in the urn. Now you draw again, what is the probability that this marble is white? Repeat the experiment where you draw a marble, which you show to your friend without you seeing it, after that you just put the marble back in the urn. Now you draw a new marble that shows to be white. What is the probability that your friend saw a white marble? Are the events first marble is white and second marble is white this time dependent or independent? As expected and described in the a priori analysis (section 6.3.2), none of the students found it difficult to solve part (a) of the problem. In part (b), there was a large division in the students' solution methods. There were groups where all students were sure that the result have to be $\frac{1}{3}$, which is of course the correct answer. Then there were groups where all the students were sure that the probability was $\frac{1}{2}$, with the expected arguments described in the a priori analysis (section 6.3.2). And then there were groups where there were both students who believed the result was $\frac{1}{3}$ and $\frac{1}{2}$. The observed group was one of the groups where everyone agreed on the correct answer from the start. One of the group's students summarized their answers as follows:

> Student 6: *That's 1 out of 3. I guess the marble in his pocket can only be one of the 3 different marbles left when we know it's not the one he drew... and only one of them can be white.*
> (Translated by the author.)

However, the group disagrees a bit about whether the events are independent or not. One of the group members believes that the events must be independent, as the probabilities are equal, which is an unexpected fallacy. Presumably it derives from the definition of independence that the students saw at the end of the previous exercise. This student may remember that there was to probabilities in the definition there were supposed to be equal if the independence should hold, and then she had mistakenly swap $P(A|B) = P(A)$ by $P(A|B) = P(B|A)$, which is not true. However, she becomes convinced that it is not correct when one of the other students in the group argues the following:

> Student 7: *No, they will depend, because if the first marble is white, then the probability that the second marble is white is of course smaller. So they are probably dependent. In my opinion. I'm not really sure what dependence is either.*
> (Translated by the author.)

It is interesting that student 7 here shows that he have acquired some knowledge about the concept of independence, but still do not feel familiar with this knowledge. This causes that Student 7 clearly is unsure of his own knowledge. This may be because the students have only just (at the end of the second exercise) been introduced to independence. The students has personalized the knowledge of the concept, but since they have not actively worked with it yet, he lacks the confidence to use his new knowledge actively.

There was one of the other groups in particular where the members of the group strongly disagreed about the solution. The group was not observed, so there is no concrete data on their discussions and arguments, but the teacher overheard most of their discussion, which he passed on to the author of this thesis afterwards. The discussion was mainly between student 3 and 4, where student 3 was sure that the answer have to be $\frac{1}{2}$, since there were two white and two black marbles when the first marble was drawn. Student 4 was sure that the answer was $\frac{1}{3}$, and then tried to convince student 3 that it was the correct answer. Student 4 had first tried explanations similar to those used by the observed group. However, it did not have the convincing effect student 4 had hoped for. After several trials with different formulations, student 4 had changed strategy. He had written down the entire sample space, after which he could cross out all the outcomes where the second marble was black. The teacher said that it is an ability the student has previously learned and practiced in mathematics lessons. This means that the student used old knowledge to personalize the new knowledge by connecting the old knowledge to the new. Student 4 then showed it to student 3 while explaining what he had done. He then counted how many outcomes were left (6) and how many of those were two white marbles (2). In this way, student 3 ended up being convinced that the probability must be $\frac{1}{3}$. That a student should come up with this solution was surprising, and probably a result of the desire to convince the other student, since the student started with a more informal argument. If the students had not disagreed (e.g. like in the observed group) it would most likely have focused on their result and perhaps a more or less informal argument. In terms of the students discussion, they have felt compelled to make as much of their personal knowledge as official as possible, in order to come up with the strongest argument and thus convince the other person that they were right. This process has expanded both students' personal knowledge.

In the validation phase of part (b), it could be seen that there was a fifty-fifty division of the groups into whether they thought it was $\frac{1}{3}$ or $\frac{1}{2}$. The teacher first let a group that had answered $\frac{1}{3}$ present their argumentation (here the observed group was chosen). Then a group with the solution $\frac{1}{2}$ was allowed to present. Both groups came up with explanations similar to those given in the a priori analysis (section 6.3.2). Finally, student 4 was allowed to present his argument to the whole class. In the institutionalization phase, the teacher took student 4's argument as a point of departure, and explained how this brute force method is always safe with symmetric probability spaces, but that they could be a bit slow, especially if you had to do it by hand. The teacher then calculated the conditional probability using the definition. $P(A|B) = \frac{P(A\cap B)}{P(B)} = \frac{\frac{2}{12}}{\frac{1}{2}} = \frac{1}{3}$, where A is the event that the first marble is white, and B is the event that the second marble is white. $P(A \cap B)$ was determined by the teacher using student 4's sample space written on the board. It was also intended that the teacher should have reviewed Bayes' theorem during this institutionalization phase, but he did not manage to do so due to lack of time.

8.5 Fourth Exercise - The Monty Hall Problem

In this exercise, the students had to work on the following problem: Suppose you are on a game show and are given the choice to select one of three doors. Behind one door there is a car and behind each of the other two doors there is a goat. Once you select a door, say No 1 (which is closed), the host, who knows what is behind each door, opens another door (say No. 3), which contains a goat. You are now given the option of changing your selection to door No. 2 or sticking with door No. 1. What would you do? (Batanero, Contreras, et al. 2014, pp. 366).

As previously mentioned, this exercise turned out quite differently than expected, as there was a very large part of the class who knew about the Monty Hall problem in advance. The teacher was well aware of this, and therefore asked the students who knew about the problem in advance not to reveal anything to the others, and to play along, which easily can be challenging, especially if students only knew the answer in advance and not the reasoning. Since there was at least one student who knew about the problem in advance in most of the groups, there was a convergence in the groups' solutions. It did not seem like the students who knew the problem revealed anything to the others, but when the other students quickly sought out the students who knew the argument and the solution for confirmation every time they made a claim. In this way, the incorrect solutions were screened out. This could be an effect of the didactic contract (although the hope was that it would fade into the background), as the students want to answer the question correctly to such an extent that they need confirmation of their answer before they make their conclusion.

There was one of the groups where none of the students knew about the Monty Hall problem beforehand. This group was coincidentally the group that were observed. In this group there was one of the students who very quickly found the simulation function on the website, MathWarehouse 2014, in the first part of their work with the problem, where they have to do manual simulations, and he started to simulate the game 1500 times. He then told the others in his group what the simulation showed and where they could find it. This convinced the other students from the group that there had to be a $\frac{2}{3}$ chance of winning if you switch doors, and $\frac{1}{3}$ if you stick. They do this solely on the basis of the simulation's output. When the students have sat and looked at the simulation for a while, one of them asks how it can be true, to which the student, who quickly found the simulation function, replies as follows:

> Student 8: *I do not know why? But it is easy to see that if you change your choice, there must be a 2/6 chance of getting the car. And if you keep your choice, then there is only 1/3 to get the car... No sorry, 2/3 not 2/6 before.*
> (Translated by the author.)

The students was surprised by this result. It seemed strange to them (as it was against their intuition), but they held on to the belief from the simulation, which is consistent with the theory (see section 4.6). This may be due to the didactic contract. It was the teacher who had handed them the simulation tool, and through the teacher's scientific authority, the website, and thus the simulation tool, have to be blueprinted as being true.

In the validation phase, it turned out that several of the groups had found the simulation function on the website, but most other groups (than the one observed) had tried different arguments than "it is true because the simulation told us". All these groups tried to argue in favor of changing the door. However, it was not easy to make a precise argument for all groups. One group wrote the following on the board:

> **We will switch.**
> 1st attempt: There is a $\frac{1}{3}$ chance of choosing a car.
> Then a door is opened with a goat. Which means that you can either choose a car or a goat.
> 2nd attempt: There is therefore a 50% chance of choosing the car, and therefore you should switch doors.

The teacher let this group explain themselves in the validation phase. The group explained that since there is only a $\frac{1}{3}$ chance of getting the car on the first choice, and that after a door has been opened, there must be a greater chance of getting the car if one made a new choice than if you refrained from making a new choice. Therefore, the group's argumentation focused more on the act of making a choice than on changing the door or not, and they forgot that not making a choice is also a choice. This means that this group basically makes the classic fallacy, described in the a priori analysis (see section 6.4.2update). The rest of the groups all had explanations close to the following reel example from, what one of the groups wrote on the board

> If you choose the strategy of switching doors and hit a goat in your first choice, you will always win a car. There is a $\frac{2}{3}$ chance of getting a goat in the first pick. There is therefore a $\frac{2}{3}$ chance of winning a car if you always switch doors.

The fact that there are so many of the groups that come up with a correct, but informal answer, is a sign that the students have expanded their personal knowledge of conditional probability. It must be remembered that there was a part of the students who knew the problem and the rationale for its results in advance. These students had expanded their personal knowledge before the lesson, and it is not certain that the lesson has made their personal knowledge greater, but perhaps more solid or easier to transform into official knowledge.

In the institutionalization phase, the teacher showed the students a table he had made beforehand that illustrated the informal explanation several of the groups had given. The table looked like this:

Door 1	**Door 2**	**Door 2**	**Stick**	**Switch**
Car	Goat - Opening	Goat	Car	Goat
Car	Goat	Goat - Opening	Car	Goat
Goat	Car	Goat - Opening	Goat	Car
Goat	Goat - Opening	Car	Goat	Car
Goat	Goat - Opening	Car	Goat	Car
Goat	Car	Goat - Opening	Goat	Car

(Translated by the author)

The teacher tells the students that he assumes that door 1 is always chosen, then all the rows will be equally likely, and thus there is a $\frac{2}{3}$ chance of winning if the door is switched. After this, the teacher presented a theorem to the students While the teacher wrote the theorem on the board, he explained to the students what a partition is. This is what the teacher wrote on the board:

> If the events, H_1, H_2,..., H_n is a partition of the sample space, S, and $P(H_i) \neq 0$ for all i, then the following holds for every event A in S:
>
> $$P(A) = P(A|H_1) \cdot P(H_1) + \cdots + P(A|H_n) \cdot P(H_n)$$

The teacher then formulated the Monty Hall problem by letting D_i be the event that the car is behind door i and W be the event of winning the car. The teacher then explained to the students that by using the theorem the following holds:

$$P(W) = P(W|D_1) \cdot P(D_1) + P(W|D_2, D_3) \cdot P(D_2, D_3)$$

The teacher assumed that door 1 is always chosen. Thus, the probability of winning if you stick had to be:

$$P(W) = 1 \cdot \frac{1}{3} + 0 \cdot \frac{2}{3} = \frac{1}{3}$$

and the probability of winning if you switch:

$$P(W) = 0 \cdot \frac{1}{3} + 1 \cdot \frac{2}{3} = \frac{2}{3}$$

8.6 Fifth Exercise - The Drawer Problem

In this exercise, the students had to work on the following problem: You have three drawers with exactly two sections in each. In one of the drawers, there is a silver coin in each of the sections. In another of the drawers, there is a gold coin in each of the sections. In the last drawer there is a silver coin in one section and a gold coin in the other.

a You choose a random section in a random drawer. What is the probability that there is a gold coin in the section you have chosen? Now you open the section and see that there is actually a gold coin in the section. What is the probability that there is a gold coin in the second section of the drawer, too?

b In the example from the first part, are the probabilities that there is a gold coin in the section you choose first and that there is a gold coin in the second section of the drawer you chose a section independent?

In order to solve this problem formally, Bayes' theorem must be used. As previously mentioned, the teacher did not manage to show the students Bayes' theorem at the end of the third exercise, as was otherwise intended. Therefore, the teacher included an introduction to Bayes' theorem in the devolution. He did it this way:

> **Bayes' Theorem:**
> For events, A and B, where $P(A) \neq 0$ and $P(B) \neq 0$, it holds
>
> $$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

The problem is a counter-intuitive problem dealing with conditional probability, where the method to solve the problem lies in analyzing the problem correctly. However, this problem is not nearly as well known as the problem in the previous exercise. Therefore, none of the students knew the solution to this problem in advance. In the validation phase, it became clear that all the students had difficulties solving this problem. All groups believed that the probability that there was a gold coin in the second section of the drawer where there had been a gold coin in the first section must be $\frac{1}{2}$, since the sample space had been reduced to the possibilities of the drawer with one silver coin and one gold coin and the drawer with two gold coins, and since it would only be a success if you had hit the drawer with two gold coins. Since all the groups came to roughly the same incorrect conclusion with roughly the same arguments, the observed group probably represents how most groups work in this exercise reasonably well. This group was focused on narrowing down the sample space (like the teacher had explained to them in the institutionalization in the first exercise) by excluding the drawer with two silver coins. The group articulated it with a reference to what they have previously learned in connection with previous exercises.

> Student 9: *Because that's that with the sample space is being reduced to... Now there can only be two of the drawers left.*
> (Translated by the author.)

It can be seen that the student is trying to approach a form of formalization. It is the first time it has been observed in class, and it suggests that the importance of formalization is beginning to sink in with the students. Unfortunately for the student, he sees the sample space as the three drawers, which are then reduced to two drawers, instead of thinking of the sections as the sample space.

Although you can tell from the students in the group that they think they must have the correct conclusion, they all have doubts about it. They have doubts about it because they think it was far too easy to solve the task and even to argue for it. This is again a sign that the didactic contract has not completely faded into the background, as students do not expect the teacher to set them problems that do not match their level or the time they have been given to work on the problem. Therefore, students continue to be critical of their arguments and conclusions.

Student 9: *So there are two left. But it won't... Is it... 50%?*
Student 10: *It sounds too simple.*
(Translated by the author.)

The students become so doubtful that they end up asking a student from another group, probably because they were not allowed to ask their teacher for help, due to the framework of the teaching. This student explains to the group that in their group they disagree whether the conclusion is $\frac{1}{3}$ or $\frac{1}{2}$. He himself leans towards $\frac{1}{2}$ and he had not understood the arguments of $\frac{1}{3}$, because when he tries to explain it, he gets stuck and ends up saying "*So I don't know... I don't understand.*" (Translated by the author). This other group's discussion could have been interesting, but unfortunately it was not observed. When the group had to write their conclusion on the board, they wrote, as mentioned, like all the other groups $\frac{1}{2}$, so something suggests that no one in the group has been able to fully argue for the correct conclusion.

In the institutionalization phase, the teacher explains to the students that he will solve the problem formally using Bayes' theorem (which is still on the board). The teacher starts with the first part of the problem, where he lets A be the event that there is a gold coin in the chosen section. Here the teacher says that there are six sections in total and there is a gold coin in three of the sections, $P(A) = \frac{3}{6} = \frac{1}{2}$. After this he moves on to the second part. He tells the students that what they have to determine is $P(H_1|A)$, where H_1 is the event that the chosen section is in the drawer with two gold coins. In addition, he tells the students that $P(H_1) = \frac{2}{6} = \frac{1}{3}$ since there are two sections in the drawer with two gold coins in them, and there are still six sections in total. After this, the teacher asks if all the students agree with these statements. All the students agreed. Then the teacher wrote the problem into Bayes' theorem.

$$P(H_1|A) = \frac{P(A|H_1) \cdot P(H_1)}{P(A)}$$

The teacher asks the students what was missing before the task could be solved.

One of the students answered that $P(H_1) = \frac{1}{3}$ and $P(A) = \frac{1}{2}$ so they only needed $P(A|H_1)$. The teacher asked if any of the students knew what this probability might be. No one answered that. The teacher explained what the notation meant by saying "(...) *so we have to find the probability that the section we have chosen contains a gold coin, when we know that the section is in the drawer that contains two gold coins*". After this, many of the students could answer that the probability must be $P(A|H_1) = 1$ This is indubitably an example of the Topaze effect (which is described in section 5), as the teacher does not give the students time to think about their answers, but instead lowers the difficulty of the question so much that it is almost trivial for the students. After that, a student was allowed to calculate the probability on the board

$$P(H_1|A) = \frac{P(A|H_1) \cdot P(H_1)}{P(A)} = \frac{1 \cdot \frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}$$

This was expectedly surprising for most students. The teacher explained the result informally by explaining that there was both the possibility of choosing one and the other section in the drawer with two gold coins.

8.7 Sixth Exercise - The Three-Event Problem

In this exercise, the students had to work on the following problem: You have two fair coins, one 1 krone and one 2 krone. You want to play heads or tails with both coins at the same time, but before that you have to consider the probabilities of the following three events: the 1 krone becomes a head, the 2 krone becomes a head and exactly one of the coins becomes a head. That is, neither more nor less than one. After these initial considerations, you have to decide whether all three events are independent or dependent.

In the a priori analysis (section 6.6.2), it was described how the assumption would be that some students would try to use the formulas they have been shown through the institutionalizations in the course. However, this was not the case. There was not a single student who tried to solve the problem in a formal way, as the linguistic way was still dominant for the students. This is a clear sign that the students have not become familiar with the formulas, and that they are not a regular part of their toolbox for solving probability problems. In addition, all groups showed a poor understanding of independence, which is a sign that the students have not had their personal knowledge of independence expanded, even though it was one of the general knowledge goals of the course.

In this exercise, it became clear that the students did not have one of the most essential parts of the concept of independence conceptualized yet, namely that independence is a property that describes the relationship between two events. Independence is not a property that a single event can have independently of other events. Four out of five groups wrote something like this on the board:

> **The 1 krone turns head:** Probability: $\frac{1}{2}$. **Independent**.
> **The 2 krone turns head:** Probability: $\frac{1}{2}$. **Independent**.
> **Exactly one of the coins turns head:** Probability: $\frac{1}{2}$. **Dependent**.
> (Translated by the author.)

In several of the previous exercises, the students worked with independence, where they both heard the teacher use the concept of independence as a characteristic of the relationship between two events (including the definition of independence in more than one version), but the students themselves also used the concept correctly several times during the course. This emphasizes that students have not become familiarized with the knowledge about independence, as students change the use of the term depending on the task. In the previous problems, where the students have used the concept correctly, they only dealt with two events. This time they have to deal with three events, which could be one of the causes for this sudden lack of knowledge. In other words, the context is different, which according to section 4.4 often causes that students to have difficulties with analyzing probability problems.

It always applies that if you know two of the events, then the unknown can be determined. There were many students who could see this for the situation where they know the first two events (i.e. the outcome of the two tossed coins), but not for the other two situations where they knew the outcome of the event, that there was precisely one of the coins that had become a heads, as well as one of the other two events. The students mixed the two expected solution methods from the a priori analysis (section 6.6.2), by concluding on one hand that the first two events are pairwise independent, as it is old knowledge for students that two coin tosses are independent, on the other hand they conclude that the third event is mutually dependent of the other two events, which is in a way correct, so it may just be the terminology that is flawed.

This is backed up by the last group's more detailed conclusion. The last group wrote the following on the board:

> **The 1 krone turns head:** Probability $\frac{1}{2}$.
> **The 2 krone turns head:** Probability $\frac{1}{2}$.
> These are independent of each other, since one coin do not dependent of another coin.
> **Exactly one of the coins turns head:** is dependent of the 1 krone, since if the 1 krone turns head, then the 2 krone have to be tail or the other way around.
> (Translated by the author.)

Although the groups' argumentation appears to be completely wrong and examples of false reasoning, it is not certain that it is so bad. The students clearly have linguistic and terminological challenges, but hidden behind this, there may be a correct conceptualisation.

Two of the students in the observed group have at the end of the action and formulation phase a discussion, that could have led them to the right final conclusion, namely that the events were mutually dependent. One student tried to convince the other that the outcome of the last event could be determined from the outcomes of the first two, which he succeed. But as it can be seen below, he came close to being able to convince himself and the other student that this also applied to the other situations.

> Student 11: *But if we think about it. If there is to be one tail and one head. And the 1 krone is a tail. Then the other will have to be a head. Then it cannot become a tail. It have to mean something... They have to dependent of each other.*
> (Translated by the author.)

The group, with these two students, was one of the four groups that answered a bit undetailed, and thus student 11 did not succeed in completing his argument.

All the groups were uncertain about their conclusions, which could also be seen in one of the groups writing on the board. Here, that group had written "*definitely wrong*" under their conclusion. This could of course be because the students thought it was a difficult exercise (which would be understandable since they do not have old knowledge of independence of three or more events) and they had no good argument that their conclusion was correct, but it could also be due to, that students no longer trusted their intuition, and that this led to an uncertainty about their own abilities.

In the observed group, the uncertainty was also clear. There was often one of the members who exclaimed "*I don't know*". The students did not trust even their most trivial argumentation, which e.g. can be seen here:

> Student 12: *You can say that the first two events must have a probability of 1/2. It's straightforward.*
> Student 13: *We say that now, but in a little while... Then it's just completely wrong.*
> (Translated by the author.)

The students have simply become accustomed to the fact that the exercises are counter-intuitive, and that their first impulse therefore is incorrect. This means that the didactic contract has conformed to the use of counter-intuitive problems, so that students now have an expectation that the teacher will give them problems where their first guess always will be wrong. This was in no way intended for the use of these problems. Unfortunately, this can lead students to think that the teacher has simply constructed some unnatural problems that are in no way relevant to their further contact with probability theory, as the problems may feel like special problems under special circumstances.

Below is a comment from a student to the teacher.

> Student 14 to the teacher: *You have nevertheless managed to turn something that should seem incredibly simple into something that requires an incredible amount of thought.*
> (Translated by the author.)

Student 14 is one of the students in the class who are confident in their own mathematical abilities, which meant that the least self-confident students properly did not dare to answer the counter-intuitive problems because they felt, that their intuition was suddenly useless, while the more confident students, as a minimum, spent much longer on even the simplest tasks, as they wanted to be completely on the safe side.

In the institutionalization phase, the teacher first reviewed the definitions of independence again to remind the students of the terminology. He then asked the students whether they were most inclined to believe that the events were dependent or independent. The students did not have to justify their answers, but simply give their answers by show of hands. There was a majority of students who believed that the events must be independent.

The teacher then made an argument for each of the claims by formally showing that the events are (pairwise) independent.
A: The 1-krone becomes heads; B: The 2-krone becomes heads; C: Exactly one of the coins becomes heads.

$$P(A) = \frac{1}{2} \quad , \quad P(B) = \frac{1}{2} \quad , \quad P(C) = \frac{2}{4} = \frac{1}{2}$$
$$P(A \cap B) = \frac{1}{4} \quad , \quad P(A \cap C) = \frac{1}{4} \quad , \quad P(B \cap C) = \frac{1}{4}$$

Which means that

$$P(A) \cdot P(B) = P(A \cap B) \quad , \quad P(A) \cdot P(C) = P(A \cap C) \quad , \quad P(B) \cdot P(C) = P(B \cap C)$$

The teacher then argued that the events were dependent by explaining to the students that you could always determine the outcome of one of the events if you knew the outcome of the other two. He did this by going over how to do it for all three situations. The teacher once again had the students vote on which conclusion they believed the most. This time the students were evenly distributed. Finally, the teacher explained to the students that it was the last conclusion that was the correct one, and then wrote the definition for mutually independence on the board, after which he showed the students the formal argument that the events were mutually independent.

8.8 Comprehension Tests

As mentioned in section 7, two short tests of approximately 10 minutes duration were carried out. The first before the teaching started and the second after the teaching had finished. If you wish to read more about the tests, please go to section 7 or if you want to see the tests go to appendix A.8 to A.11, where the tests can be found in both the original version and a version translated into English. Normally there are 27 students in the class. At the first test, there were a total of 19 students present, all of whom took part in the test. At the second test, a total of 16 students were present, all of whom took part in the test. There were 13 of the students who took part in both the first and the second test.

In the first part of the tests, students had to answer two general comprehension questions. They were asked the same questions on both tests. In the first question in the first test, there were 4 out of the 19 students, who did not answer anything that was conceptually incorrect, but their answers were not very profound either. All 4 of them answered something close to "*Something that does not depend on anything else*", which is a quote from one of them. This is a linguistic approach, and an expected approach, since the students did not know the concept of independence beforehand. It is surprising that there were not more of the students who gave an answer of this type, since it was expected that the linguistic approach would be the most immediate for these students. What is special about these four students is that none of them did particularly well in the second part of the test. In this part, students are given three tasks where they have to decide whether two events are independent or not. The four students had only one of these tasks correct in total. This means that three out of four of the students did not answer any of the tasks in the second part correctly. In comparison, there was only one other student out of the remaining 15 students who did not answer any of the tasks in the second part of the test correctly. This means that this faint and inexact explanation of independence is a marker of students without a clear and fruitful intuitive conceptualization.

Over half of the students (10) at the first test explain how independence is when a variable is not dependent and therefore not affected by another variable. This suggests that the students draw on old knowledge from another field of mathematics, namely functional theory. In the second part of the test, these students had an average of 1.8 out of 3 correct answers. This suggests that even if these students answer something that is conceptually wrong, their intuitive conceptualization is still more useful than the four who had given a more correct answer in the first question. However, one must be cautious in concluding something based on these tests, as it is a small amount of data. Of the 10 students with the functional theory approach, there were 6 of them who took part in both tests. Two of the six students largely did not change explanations of independence in the second test. However, it was only one of the others who changed the description of independence to something more correct. That student wrote in the second test:

> *That events are independent means that they do not affect each other. Eg. if I roll with a die, the independence will consist in the fact that the first roll has no effect on what I roll in the next roll. Conversely, if events are dependent, this means that one event has an impact on another event.* (Translated by the author.)

It is clear that this student has made progress, not only in language use, but also in mathematical understanding of independence. In the following question, this student also explains how two events are independent if the formula $P(A \cap B) = P(A)P(B)$ is true. Unfortunately, this is not the case for the other three students. They all try to explain why something must be equal to each other (just like in the definition the student from before could remember). However, none of them can remember the formula. All three explain themselves pretty much like this example (there is a quote from one of them):

> *When you have two events, you can calculate whether they are independent by seeing if there is the same probability on both sides of the equals sign, and if there is, they are independent.*
> (Translated by the author.)

It is quite normal that students are not used to memorizing formulas, as they are used to having their collection of formulas with them. It is not necessarily a success criterion that the students can remember formulas, but since one of the goals of the teaching design has been that the students should train the use of formalization in probability-theoretic problem solving, it is important to investigate whether the students have become better at using formulas. It is clear that these three students have the definition of independence in mind, which is a sign of an increased focus on formalization. In addition, the students are used to doing many type tasks where they have to use the formulas they have to learn to use. In this course, students have not done nearly as many tasks as they are used to or tasks with the same degree of uniformity.

In the second part of the tests, a general analysis of the students' skills on independence was made easily, by simply looking at the amount of correct answers to each of the three tasks. The first task in the second part of the test requires the students to decide whether two events (they are different in the two tests) are independent or not, when they know that $P(A) = 10\%$ and $P(A|B) = 10\%$ in the first test, and that $P(A) = 14\%$ and $P(A|B) = 14\%$ in the second test, respectively. In this task, 42.1% of the students answered that they are independent (which is of course correct), while the rest answered that they are dependent. In the second test, 62.5% answered "independent" (which is still correct), while the rest answered "dependent". Unfortunately, not many gave reasons for their answers, but there were a few who did. Here they explained that, in the first of the tests, they judged the events to be either independent or dependent, as they either believed or did not believe that the gender of a car's driver would have a correlation with the color of the car, which is an approach of a subjectivist (see section 4.1). In the last of the tests, there were only two who wrote their reasons. One still made the same kind of subjective judgment, but the other wrote; "*From the probabilities, you can see that nothing happens. Therefore, they have to be independent.*" This approaches an objectivist way of thinking as it has elements of formalization. The increase in the percentage of correct answers is probably due to the fact that there are more of the students who have trained their abilities of formalization and thereby conceptualized a correct knowledge of independence.

In the second task in the second part of the tests, the students had to decide whether two events were independent or not, where they were given the probabilities of $P(A \cap B)$, $P(A)$ and $P(B)$. In this task, there were 68.4% of the students who answered correctly on the first test, while there were 81.3% who answered correctly on the second test. It is assumed (as there are no arguments for many of the students' answers), that many of the students made a subjective judgment in the first of the tests. The reason that there is a larger percentage who answer correctly in the second of the tests is difficult to deduce. It could be because the students had trained their formalization and could remember the definition of independence, but it could just as well be because the task was worded differently or purely coincidental. The reason why the percentage of correct answers on both tests is so high is probably that students tend to judge (subjectivist) that the events are dependent (which is the correct answer) when making their personal, not mathematical, judgment.

In the last of the three tasks in the last part of the tests, only 21.1% of the students answered correctly in the first test, while 31.3% of the students answered correctly in the second test. These two tasks were designed so that students who made subjectivist judgments would not answer correctly. The task in both the first and the second test's version is done in the same way as the first of the tasks, so one could decide whether the events are independent or not by using the definition of independence that uses conditional probability. This time the events are just dependent instead of independent. However, the events are selected so that most people will think that they are independent, if they use a subjectivist approach. This could indicate that there are few students in both the first and second test who do not use subjectivist judgment. This must be taken with a slight caveat that there may be students who do not exclusively use this form of judgment, but approches the problem in a way that is a mixture of subjectivist and objectivist. Since a personal judgment will clearly state that the events are independent, the students' subjectivist intuition will therefore be clear about the independence. This may mean that even if they are able to use the correct definition of independence, their belief in this definition that they have just learned is not necessarily strong enough to overcome the students' intuition.

8.9 The Teachers and the Students' opinion

In the methodology section (section 7) it was described how an interview with the teacher after the course should be part of the data collection. This was carried out as planned. Since the course took a bit less time than expected, there was a little time to spare. The author of the thesis chose to use this to create a plenary discussion, where the students could talk about their experiences with and opinions on the teaching. It must be said that the teacher was in the classroom while this discussion was taking place, and it was before the interview with the teacher. Therefore, there is a possibility that the teacher may have been influenced by the students' input.

The students and the teacher had especially one opinion on the teaching. They felt that a more repetitive part was missing, where students could work with several more or less monotonous tasks, where they could use the relevant formulas several times and thus become more experienced and familiar with the formulas and working methods. The students found it frustrating to work with what they thought were isolated exercises where they had to do something new each time. The students also said that it had been exciting to try to work in a different way than they were used to in their mathematics lessons. They particularly liked the times when they had disagreed about solutions and through discussions were able to conclude what was correct and what was incorrect. On the other hand, they found it very frustrating that they could not rely on their intuition in the exercises. One of the students even said:

> *It was honestly maddening that everything we thought was right was wrong. It has really annoyed me. But then I also have to say that I can remember the exercises we did. This is not always the case in mathematics.*
> (Translated by the author.)

Then the question is whether it was only on the basis of irritation that the exercises had stuck in the student's memory or whether the surprise that his intuition was not correct had perhaps also made the exercises memorable.

In addition to the lack of routine-creating tasks, the teacher also had the frustration that he did not think that the learning outcome was worth the time spent. The teacher himself said that he knew that if the students and himself became more accustomed to teaching in this way it would perhaps become more effective, but he was not convinced that it could become effective enough in respect to the relationship between the amount of subjects in the curriculum and the number of teaching hours. The teacher was pleasantly surprised by the formulation phases, where the students had to write conclusions on the board. He had had the experience that through these phases the students were forced to sharpen their mathematical language and thus become more precise in their formulations. In addition, he thinks it helped the students to finish their work properly, and not just think a solution halfway through and then stick to it. The teacher thought the validation phases had been difficult. He could sense that they were merging with the institutionalization phases. He believed that this was because he was very used to teaching in a way that resembled a mixture of validation and institutionalization. However, he had the feeling that the institutionalization filled more than the validation, and that it was because he had felt that there was a lot he had to go through in the institutionalization phases.

9 Discussion

The first and biggest critique of this study is that it is difficult to conclude anything from such a narrow data base. Only one single class took part in the study, and since there may be special circumstances with this class, developments or lack of the same could be due to the class' special circumstances as to the effects of the prepared teaching design. A larger study with several participating classes from many different places in Denmark and with different fields of study could have contributed to a much stronger conclusion, where the special circumstances of the classes would fade away. This could not be done in this thesis, as there is extremely limited time (only four months) to prepare the thesis.

One of the teacher's comments on the teaching design (see section 8.9) was a concern about the relationship between time consumption and relevance in relation to the curriculum. Is it worth spending a minimum of five lessons working with conditional probability and independence, now that it is not directly part of the curriculum? One of the arguments for spending time on a course, which is described in this thesis, is that it should give the students a good and well-founded foundation, on which they can build solid knowledge in relation to other subjects that are in the syllabus. Therefore, it could have been interesting to investigate how the students' conceptualization of conditional probability and independence might have been a strong foundation for other topics. The best example in upper secondary school is binomial distributions, where the students, as mentioned in section 2, must use independence as a basic condition for their theory. More specifically, it would have been interesting to investigate whether students with a more complete conceptualization of independence had better opportunities to conceptualize the theory of binomial distributions than students who have not worked in the same way with independence. This was again not an option due to the limited time in which the thesis had to be done, which meant that there was no time to return to the class and investigate the relationships with binomial distributions later in the process, but also because the investigation itself was scheduled for five lessons, and thus also limited.

In section 8.9, it was explained how both the teacher and the students think routine-creating type tasks were missing in the teaching design. It was a conscious choice not to include type tasks, as one of the main ideas behind the design was that the students should train their analytical skills on probability problems, i.e. the part of problem solving where you find out what the problem is about and thus how to solve it. Type tasks are tasks where you get the students used to working with problems where exactly the same analysis must be done each time. This means that the analysis part takes a back seat, as it is the same for all tasks, and the students can therefore focus on working with their calculation skills and thus turn them into routines. Thus, routine teaching will reduce students' training in analyzing problems. The reason why including routine type tasks can be justified is that they can strengthen students' use of formalism. In this project, it was seen several times that students did not reach a phase of problem solving where they could use formalization because they could not analyze the problem correctly. In addition, students can train their formalization skills much more effectively with type tasks, as they are not as time-consuming as these larger problems used in this thesis. In order to accommodate both the students' analytical abilities and their formalistic abilities, it could be fruitful that the design was a mixture of the more time-consuming problems, where the students can work on their analytical abilities, and of the routine type tasks, where the students could develop their formalistic abilities. If it were to be introduced in the teaching design of this thesis, it would require either one or two of the exercises to be taken out in favor of routine tasks or that the teaching course be extended. This applies despite the fact that the course ended up being around 20-25 minutes shorter than expected.

Another focal point in the teaching design was to strengthen the students' linguistic abilities within probability theory. When looking at the comprehension tests, it is clear that the students use a more probability-theoretic language after the teaching course than before. There were many students who started by talking about independence of variables, and a few who wrote about independence of outcomes and/or information. By the end of the course, it had mostly been replaced by independence of events. This suggests that the students have strengthened their conceptualization of independence. During the action phases of the exercises, it could be observed that the students' use of language also became more probabilistic, or in other words the students moved into a probabilistic terminology. Once again, this may be due to students developing their conceptual understanding of independence and conditional probability, but it may also be an example of the Jourdain effect. It may be that the students have simply adopted the linguistics from the teacher's use of language, without having gained an increased understanding of the meaning of the words. In the first exercise, it could be seen how one student developed her linguistic precision and thus suddenly understood connections in the task that she had not been able to before (see section 8.2 for a deeper review). This suggests that the student has not simply adopted a language she does not understand, as she actively brings the concepts into play.

In the second part of the comprehension tests, where the students had to decide whether two events were independent or not, were not so successful, as most of the students did not write reasons for their answers. This had been far more useful than just the students' conclusions. Thus, this part was largely narrowed down to a comparison between the amount of correct answers the students could give in each of the tests. Here, a small improvement could be seen, which in principle indicates that the students have become better at using both their analytical and formalistic abilities. However, the progress is not very great, and since the tasks are slightly different from the first test to the second test, the progress could also be due to several of the students accidentally concluding correctly, but with a false reasoning. However, the fact that there is progress in all three tasks does not indicate this, as the students could just as well have had a decline in a single one of the tasks. However, since there are only three tasks, this is difficult to conclude, since there would be a 12.5% chance that there was a progress in all three tasks, assuming that it is equally likely to have a progress or a decline in the question based on the wording of the questions.

10 Conclusion

The overall research questions of this thesis were as follows

> *How should conditional probability and independence be taught so that students acquire solid knowledge about the subject that is conceptualized in a non-misleading way?*

In an attempt to answer this question, this thesis has investigated the way in which conditional probability and independence are included in curricula and textbooks. It can be concluded that the subject is no longer part of upper secondary education, since conditional probability is not part of curricula or textbooks anymore, and since independence is only appearing as an assumption. In the course of the thesis, several misconceptions within conditional probability and independence have been reviewed, including the fallacy of the time axis, the base rate fallacy, the fallacy of the transposed conditional and the conjunction fallacy. In addition, it has been reviewed how i.a. contraintuitive problems, simulations, contingency tables, and linguistic training can help the students offset these fallacies.

Based on this, a teaching design has been made, which was unfolded over five lessons of 55 minutes. From observations and data from this teaching course, it can be concluded that the students have developed their linguistic abilities within probability theory, which leads to an increased possibility of being able to carry out correct analyzes of probabilistic problems. This was one of the goals of the teaching design. Another goal was for the students to develop their formalistic skills. This was only successful to a very low degree, and therefore it can be concluded that more lessons must be used to give the students the opportunity to do this, or that the five lessons must be reprioritized so that there is time for routine tasks on behalf of one or more of the counterintuitive unique problems. It is not possible, based on the collected data and observations, to be able to conclude that the teaching design reject misleading conceptions, and that would thus require a more comprehensive teaching design both in quality and quantity. In addition, it would also require a larger study with more participating subjects.

References

Bar-Hillel, Maya (1983). "The base rate fallacy controversy". In: *Decision making under uncertainty*. Ed. by Roland W. Scholz. Amsterdam, The Netherlands: Elsevier, pp. 39–61.

Batanero, Carmen, J. Miguel Contreras, Carmen Díaz, and Gustavo R. Cañadas (2014). "Preparing Teachers to Teach Conditional Probability: A Didactic Situation Based on the Monty Hall Problem". In: *Mit Werkzeugen Mathematik und Stochastik lernen = Using Tools for Learning Mathematics and Statistics*. Ed. by Thomas Wassong, Daniel Frischemeier, Pascal R. Fischer, Reinhard Hochmuth, and Peter Bender. Weisbaden, Germany: Springer. Chap. 26, pp. 363–376.

Batanero, Carmen, Antonio Estepa, Juan D. Godino, and David R. Green (Mar. 1996). "Intuitive Strategies and Preconceptions about Association in Contingency Tables". In: *Journal for Research in Mathematics Education* 27.2, pp. 151–169.

Batanero, Carmen, Michel Henry, and Bernard Parzysz (2005). "The Nature of Chance and Probability". In: *Exploring Probability in School: Challenges for Teaching and Learning*. Ed. by Graham A. Jones. New York NY, USA: Springer. Chap. 1, pp. 15–37.

Batanero, Carmen and Ernesto Sanchez (2005). "What is the Nature of High School Students' Conceptions and Misconceptions about Probability?" In: *Exploring Probability in School: Challenges for Teaching and Learning*. Ed. by Graham A. Jones. New York NY, USA: Springer. Chap. 10, pp. 241–266.

Benhamou, Eric and Valentin Melot (Sept. 2018). "Seven proofs of the Pearson Chi-squared independence test and its graphical interpretation". In: pp. 1–20.

Blitzstein, Joseph K. and Jessica Hwang (2014). *Introduction to Probability*. Boca Raton FL, USA: CRC Press.

Borovcnik, Manfred (Sept. 2012). "Multiple perspectives on the concept of conditional probability". In: *Avances de Investigación en Educación Matemática* 2, pp. 5–27.

Borovcnik, Manfred and Robert Peard (1996). "Probability". In: *International Handbook of Mathematics Education*. Ed. by Alan J. Bishop, Ken Clements, Christine Keitel, Jeremy Kilpatrick, and Colette Laborde. Dordrecht, The Netherlands: Kluwer Academic Publishers. Chap. 7, pp. 239–288.

Brousseau, Guy (1997). *Theory of Didactical Situations in Mathematics: Didactique Des Mathématiques*. Ed. by Nicolas Balacheff, Martin Cooper, Rosamund Sutherland, and Virginia Warfield. Dordrecht, The Netherlands: Kluwer.

Brydensholt, Morten and Grete Ridder Ebbesen (2011). *Lærebog i Matematik Bind 2*. Viborg, Denmark: Systime.

Carles, Marta and M. Pedro Huerta (Feb. 2007). "Conditional Probability Problems and Contexts: The Diagnostic Test Context". In: *Proceedings of the Fifth Congress of the European Society for Research in Mathematics Education, CERME* 5.2, pp. 702–710.

Carstensen, Jens, Jesper Frandsen, and Jens Studsgaard (2006). *MAT B2 stx*. Viborg, Denmark: Systime.

Clausen, Flemming, Gert Schomacker, and Jesper Tolnø (2006a). *Gyldendals Gymnasiematematik Arbejdsbog B2*. Gylling, Denmark: Gyldendal.

— (2006b). *Gyldendals Gymnasiematematik Grundbog B2*. Gylling, Denmark: Gyldendal.

Common Core State Standards for Mathematics (Oct. 2022). URL: `http://www.corestandards.org/wp-content/uploads/Math_Standards1.pdf`.

Curriculum of Norway (Oct. 2022). URL: `https://sokeresultat.udir.no/finn-lareplan.html?fltypefiltermulti=Kunnskapsl%5C%C3%5C%B8ftet%5C%202020&page=8`.

Curriculum of Sweden (Oct. 2022). URL: `https://www.skolverket.se/undervisning/gymnasieskolan/laroplan-program-och-amnen-i-gymnasieskolan/gymnasieprogrammen/amne?url=-996270488%5C%2Fsyllabuscw%5C%2Fjsp%5C%2Fsubject.htm%5C%3FsubjectCode%5C%3DMAT%5C%26version%5C%3D11%5C%26tos%5C%3Dgy&sv.url=12.5dfee44715d35a5cdfa92a3`.

Díaz, Carmen and Carmen Batanero (Oct. 2009). "University Students' Knowlegde and Biases in Conditional Probability Reasoning". In: *International Electronic Journal of Mathematical Education* 4.3, pp. 131–162.

EVA, The Danish Evaluation Institute (2009). *The subject of Mathematics from an international perspective: Mathematics A and B in HTX and STX*.

Falk, Ruma (1989). "Inference under Uncertainty via Conditional Probabilities". In: *Studies in Mathematical Education: The Teaching of Statistics*. Ed. by Robert Morris. Paris, France: Unesco. Chap. 14, pp. 175–184.

Feller, William (1968). *An introduction to probability theory and its applications*. New York, USA: Wiley.

Finetti, Bruno de (1974). *Theory of Probability*. New York NY, USA: Wiley.

Freudenthal, Hans (1973). *Mathematics as an Educational Task*. Dordrecht, The Netherlands: D. Reidel.

Gras, Régis and André Totohasina (Jan. 1995). "Chronologie et causalité, conceptions sources d'obstacles épistémologigues à la notion de probabilité conditionnelle [Chronology and causality, conceptions sources of epistemological obstacles in the notion of conditional probability]". In: *Recherches en Didactique des Mathématiques* 15.1, pp. 49–95.

Jessen, Britta and Carl Winsløw (2018). *MERIA Template for Scenarios and Modules*. Ed. by Sanja Antoliš, Jeanette Axelsen, Matija Bašic, Rogier Bos, Kristijan Cafuta, Aneta Copić, Gregor Dolinar, Michiel Doorman, Željka Milin Šipuš, Selena Praprotnik, Sonja Rajh, Mateja Sirnik, Mojca Suban, Eva Špalj, Carl Winsløw, Petra Žugec, and Vesna Županović. MERIA.

Kelly, I. W. and F. W. Zwiers (1988). "Mutually Exclusive and Independence: Unravelling Basic Misconceptions in Probability Theory". In: *The Proceedings of the Second International Conference on Teaching Statistics*. Ed. by R. Davidson and J. Swift. Victoria B.C., Canada: University of Victoria, pp. 96–100.

Kolmogorov, Andrey N. (1933/1956). *Foundations of the Theory of Probability*. 2nd ed. New York NY, USA: Chelsea.

Læreplan Matematik A - STX (Aug. 2005). URL: https://www.gymnasieforskning.dk/wp-content/uploads/2013/09/Mathematics-from-an-international-perspective.pdf.

Læreplan Matematik A - STX (June 2013). URL: https://www.retsinformation.dk/eli/lta/2013/776#Bil35.

Læreplan Matematik A - STX (Aug. 2017). URL: https://www.uvm.dk/gymnasiale-uddannelser/fag-og-laereplaner/laereplaner-2017/stx-laereplaner-2017.

MathWarehouse (2014). *Monty Hall Simulation Online*. Last accessed 01 November 2022. URL: https://www.mathwarehouse.com/monty-hall-simulation-online/.

Molnar, Robert Adam (2015). "High School Mathematics Teachers' Knowledge and Views of Conditional Probability". PhD thesis. The University of Georgia.

Shaughnessy, J. Michael (1992). "Research in Probability and Statistics: Reflection and Directions". In: *Handbook of Research on Mathematics Teaching and Learn-*

ing. Ed. by Douglas A. Grouws. Reston, VA, USA: National Council of Teachers of Mathematics, pp. 465–494.

Sørensen, Michael (2013). *En Introduktion til Sandsynlighedsregning [A Introduction to Probability]*. 13th ed. Copenhagen, Denmark: University of Copenhagen.

Székely, Gábor J. (1986). *Paradoxes in Probability Theory and Mathematical Statistics*. Dordrecht, The Netherlands: D. Reidel.

The National Curriculum of England (Dec. 2014). URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/840002/Secondary_national_curriculum_corrected_PDF.pdf.

Tversky, Amos and Daniel Kahneman (1982). "Judgements of and by representativeness". In: *Judgment under uncertainty: Heuristics and biases*. Ed. by Daniel Kahneman, Paul Slovic, and Amos Tversky. New York, USA: Cambridge University Press. Chap. 6, pp. 84–98.

Vejledning Matematik A - STX 2017 (Oct. 2021). URL: https://www.uvm.dk/gymnasiale-uddannelser/fag-og-laereplaner/laereplaner-2017/stx-laereplaner-2017.

A Appendix

A.1 Lesson Plans - Danish Version

A.1.1 First Exercise

Fase	**Lærerens rolle**	**Elevernes rolle**
Introduktion til miniforløb *2-3 min*		Lytter.
Test for indsamling af data *10-15 min*		Besvarer testen.
Devolution (didaktisk) *3-5 min*	Fremviser kontingenstabel og sikre sig, at eleverne forstår tabellens indhold. Fremviser arbejdsspørgsmål. Se under skema for at finde kontingenstabel og arbejdsspørgsmål. Fortæller eleverne at de skal arbejde med deres sidemakker. Hvert par skal have et papir med kontingenstabellen og arbejdsspørgsmål, der kan findes under dette skema.	Lytter og stiller opfølgende spørgsmål. Modtager papiret og gør klar til makkerskabet.
Handling og formulering (adidaktisk) *10-15 min*	Observerer eleverne uden at interagere med dem. Der kan dog interageres med elevpar, der ikke kan komme i gang eller er stoppet helt op og ikke selv kan komme i gang igen.	Arbejder med arbejdsspørgsmålene. Skriver deres løsninger ned på papiret.

Validering (didaktisk) *7-10 min*	Gennem lærerens observationer i handlings- og formuleringsfasen har læreren fået et overblik over, hvilke løsninger de forskellige makkerpar har. Derfor kan læreren udvælge forskellige makkerpar til at forklare deres løsningsforslag.	Lytter når læreren eller klassekammerater taler. Forklarer deres løsninger, når de bliver spurgt af læreren. Deltager aktivt ved spørgsmål i plenum.
Institutionalisering (didaktisk) *10 min*	Gennemgår betinget sandsynlighed teoretisk med udgangspunkt i de netop afsluttede opgaver. Benytter Venn-diagrammer til at illustrere sandsynligheder og indskrænkningen af udfaldsrummet ved betingelse.	Lytter og tager evt. notater.

A.1.2 Second Exercise

Fase	**Lærerens rolle**	**Elevernes rolle**
Devolution (Didaktisk) *2-3 min*	Fremviser kontingenstabel, som ses under dette skema. Forklarer at kontingenstabellen beskriver en anden test for sygdommen Divoc, som de arbejdede med i sidste lektion. Giver eleverne det tilhørende arbejdsspørgsmål; ”Er denne test for Divoc en god test? Hvorfor/hvorfor ikke?”. Inddel eleverne i grupper af 4-5.	Lytter og sætter sig derefter i deres grupper.
Handling (Adidaktisk) *7-9 min*	Observerer eleverne. Interagerer ikke med dem, med mindre, det er nødvendigt for at få en gruppe i gang med arbejdet.	Arbejder med arbejdsspørgsmålet.

Formulering (Adidaktisk) *3-4 min*	Inddeler tavlen, så der er et område til hver gruppe, og fortæller eleverne, at hver gruppe skal skrive deres løsning på tavlen. Derefter observeres eleverne.	Eleverne skriver deres løsning op på tavlen.
Validering (Didaktisk) *6-7 min*	Vælger en gruppe, der starter med at fremlægge deres løsning. Gruppen vælges på baggrund af de observationer læreren har gjort i de tidligere faser. Derefter kan denne løsning diskuteres i plenum eller der kan vælges en ny gruppe, der kan fremlægge deres løsning, hvis dette er vurderes bedre for eleverne udbytte.	Nogle eleverne fremlægger for de andre, der lytter opmærksomt. Deltager aktivt i plenumdiskussioner.
Institutionalisering (Didaktisk) *5-6 min*	Gennemgår uafhængighed teoretisk med udgangspunkt i opgaven. Kommer ind på definitioner for uafhængighed, både definitionen med og den uden brug af betinget sandsynlighed.	Lytter og stiller evt. opfølgende spørgsmål.

A.1.3 Third Exercise

Fase	**Lærerens rolle**	**Elevernes rolle**
Devolution (Didaktisk) *4-5 min*	Forklarer eleverne om urneproblemet. Eksperimentet udføres foran eleverne imens opgaven forklares. Forklarer eleverne at de skal forsætte I deres grupper.	Lytter og stiller evt. opfølgende spørgsmål.
Handling (Adidaktisk) *8-10 min*	Observerer eleverne og interagerer kun, hvis det er nødvendigt for at få eleverne i gang med at arbejde.	Arbejder med opgaven i deres grupper.

Formulering (Adidaktisk) *2-3 min*	Giver eleverne besked om, at de skal skrive deres svar på tavlen igen. Derefter observeres eleverne igen.	Skriver løsninger på tavlen.
Validering (Didaktisk) *8-10 min*	Udvælger en gruppe til at redegøre for deres svar. Derefter kan en anden gruppe vælges til at fremlægge deres svar, hvis de har en anden forklaring eller diskussionen kan foregå i plenum. Dette kommer an på lærerens observationer.	Redegør for deres løsning, hvis de bliver udvalgt af læreren, ellers lytter. Deltager aktivt i plenumdiskussioner.
Institutionalisering (Didaktisk) *3-4 min*	Forklarer eleverne om hvordan betinget sandsynlighed og kausalitet ikke er det samme. Hvordan betinget sandsynlighed kan benyttes modsat tidsforløb. Hvis der er tid kan læreren komme ind på Bayes sætning.	Lytter og stiller evt. opfølgende spørgsmål.

A.1.4 Fourth Exercise

Fase	**Lærerens rolle**	**Elevernes rolle**
Devolution (Didaktisk) *8-10 min*	Giver eleverne et ultrakort rids af historien bag Monty Hall-problemet. Forklarer eleverne spillereglerne til Monty Hall. Spil et par eksempler igennem, hvor eleverne kan være med til at udvikle strategier (f.eks. start altid med en specifik dør, skift dør eller bliv ved dør).	Lytter, deltager i plenumundervisningen og stiller evt. opfølgende spørgsmål.

	Hvis eleverne ikke selv nævner strategier om valget af den første dør, undlades denne del for effektivisering. Hvis de nævner disse strategier, er der nødvendigt at udelukke disse, hvilket gøres ved i plenum at komme frem til hvorfor valget af dørnummer ved første valg ikke har indflydelse på vinderchancen. Bed evt. eleverne om at tage stilling til, om de mener at en strategi ville være fordelagtig, og hvis de mener det; hvilken. Del link til hjemmeside. Forklar eleverne, at de nu hver for sig får 5 min til at spille Monty Hall-spillet. (Altså den del af den delte hjemmeside, hvor man skal være "quiz-deltageren".) Fortæl eleverne, at de selv må vælge om de spiller med en strategi eller på ren intuition. Hvis de spiller med en strategi, må de selvfølgelig også selv vælge, hvilken strategi de benytter. Hjemmesiden laver selv statistik. Link til hjemmesiden kan findes under dette skema, hvor der også er en forklaring af, hvor man finder statistikken.	
Handling (Adidaktisk) *6-7 min*	Observerer eleverne. Interagerer ikke med dem, med mindre, det er nødvendigt for at få en elev i gang med arbejdet.	Spiller spillet selvstændig.

Formulering (Adidaktisk) *6-7 min*	Inddeler eleverne i grupper på 4-5 elever pr. gruppe. Beder eleverne sammenligne deres statistik, og derefter at formulere en foreløbig konklusion på, hvordan man bør spille spillet for at have størst chance for at vinde, og skrive den på tavlen. Derefter observerer læreren eleverne.	Sammenligner deres statistik med hinanden, for at få den bedste baggrund for at formulere en foreløbig konklusion på, hvordan man bør spille spillet for at have størst chance for at vinde. Skriver deres foreløbige konklusion på tavlen.
Validering (Didaktisk) *8-10 min*	Får et par af grupperne til at fremlægge deres foreløbige konklusion, og hvorfor de er kommet til denne konklusion.	Fremlægger konklusioner og argumenter eller lytter, hvis de ikke fremlægger.
Devolution (Didaktisk) *2 min*	Forklarer eleverne, at de på hjemmesiden også kan simulere rigtig mange gentagelser af spillet på kort tid. Beder eleverne om ligesom før først selvstændig at arbejde med simulationen og derefter sætte sig sammen med deres grupper og diskutere deres (måske) nye konklusion, skrive den på tavlen, samt overveje, hvordan det kan forholde sig sådan. De skal også skrive deres argumenter på tavlen.	Lytter og stiller evt. opfølgende spørgsmål.
Handling (Adidaktisk) *3-5 min*	Observerer eleverne. Interagerer ikke med dem, med mindre, det er nødvendigt for at få en elev i gang med arbejdet.	Arbejder med simulationen selvstændigt.

Formulering (Adidaktisk) *10-12 min*	Observerer eleverne. Interagerer ikke med dem, med mindre, det er nødvendigt for at få en gruppe i gang med arbejdet.	Arbejder i grupper med at formulere deres konklusion på hvilken strategi, der giver den største vinderchance, skriver den på tavlen under deres tidligere konklusion (det gør ikke noget, hvis det er den samme konklusion). Diskutere internt i gruppen, hvordan det kan være, at deres konklusion passer. Skriver deres argumenter på tavlen under deres konklusioner.
Validering (Didaktisk) *7-10 min*	Lader en gruppe eller to forklarer deres konklusion og deres argumenter for resten af klassens elever. Åbner diskussionen op i plenum for andre argumenter.	Forklarer konklusioner og argumenter, lytter og deltager i klassediskussionen.
Institutionalisering (Didaktisk) *7-10 min*	Gennemgår den formelle løsning og evt. en mindre formel løsning, hvor der benyttes tælletræer.	Lytter og stiller evt. opfølgende spørgsmål.

A.1.5 Fifth Exercise

Fase	**Lærerens rolle**	**Elevernes rolle**
Devolution (Didaktisk) *2-3 min*	Forklarer kommodeproblemets første del (som kan ses under skema). Fortæller eleverne, at de først selv skal finde en løsning i 2 minutter og derefter i grupper diskuterer deres løsninger.	Lytter og stiller evt. opfølgende spørgsmål.

Handling (Adidaktisk) *2-3 min*	Observerer og klargør til inddelingen af grupper på 2-3 elever.	Arbejder selvstændigt med en løsning.
Handling og formulering (Adidaktisk) *10-12 min*	Inddeler eleverne i grupper af 2-3. Informerer eleverne om, at de skal være klar til at skrive deres løsning op på tavlen med udregninger. Fortæl dem at de har 10 min Derfra observerer læreren eleverne uden at interagerer.	Arbejder i deres grupper med at komme frem til en løsning, de er enige om. Derefter skriver de deres løsning på tavlen med argumenter.
Validering (Didaktisk) *7-9 min*	Ud fra svarene på tavlen og tidligere observationer vælges 1-3 grupper til at fremlægge deres løsningsforslag. Derefter og/eller imellem fremlæggelserne kan læreren stille spørgsmål til gruppen, der fremlægger, eller ud i plenum.	Fremlægger, lytter og/eller svarer på spørgsmål. Stiller evt. uddybende spørgsmål.
Devolution (Didaktisk) *2 min*	Giver eleverne anden del af kommodeproblemet. Problemet kan ligesom før findes under dette skema. Fortæller eleverne at processen bliver ens med før; først 2 min hver for sig og derefter 10 min i grupperne.	Lytter og stiller evt. opfølgende spørgsmål.
Handling (Adidaktisk) *2-3 min*	Observerer eleverne.	Arbejder med spørgsmålet hver for sig.
Handling og Formulering (Adidaktisk) *10-12 min*	Observerer eleverne.	Arbejder i deres grupper med at komme frem til en løsning de er enige om. Derefter skriver de deres løsning på tavlen med argumenter.

Validering (Didaktisk) *6-7 min*	Ud fra svarene på tavlen og tidligere observationer vælges 1-3 grupper til at fremlægge deres løsningsforslag. Derefter og/eller imellem fremlæggelserne kan læreren stille spørgsmål til gruppen, der fremlægger, eller ud i plenum.	Fremlægger, lytter og/eller svarer på spørgsmål. Stiller evt. uddybende spørgsmål.
Institutionalisering (Didaktisk) *7-8 min*	Forklarer de formelle løsninger til begge dele af kommode-problemet.	Lytter og stiller evt. uddybende spørgsmål.

A.1.6 Sixth Exercise

Fase	**Lærerens rolle**	**Elevernes rolle**
Devolution (Didaktisk) *3-4 min*	Forklarer eleverne om 3-hændelsesproblemet, der kan findes under dette skema. Forklarer ydermere, at de skal løse opgaven ligesom de to andre gange i denne lektion.	Lytter og stiller evt. opklarende spørgsmål.
Handling (Adidaktisk) *2-3 min*	Observerer eleverne.	Arbejder med spørgsmålet hver for sig.
Handling og Formulering (Adidaktisk) *10-15 min*	Observerer eleverne.	Arbejder i deres grupper med at komme frem til en løsning de er enige om. Derefter skriver de deres løsning på tavlen med argumenter.

Validering (Didaktisk) *8-10 min*	Ud fra svarene på tavlen og tidligere observationer vælges 1-3 grupper til at fremlægge deres løsningsforslag. Derefter og/eller imellem fremlæggelserne kan læreren stille spørgsmål til gruppen, der fremlægger, eller ud i plenum.	Fremlægger, lytter og/eller svarer på spørgsmål. Stiller evt. uddybende spørgsmål.
Institutionalisering (Didaktisk) *7-10 min*	Gennemgår den formelle forklaring på, hvorfor de 3 hændelser ikke er uafhængige, når de nu er parvist uafhængige. Trækker evt. linjer til binomialfordelinger.	Lytter og stiller evt. opfølgende spørgsmål.
Test for indsamling af data *10-15 min*		Besvarer testen.

A.2 First Exercise - The First Contingency Table (Danish Version)

Tabellen viser opdigtet data fra en stikprøve der skal hjælpe med at afklare effekten af at masseteste en befolkning for en bestem opdigtet sygdom (kunne hedde Divoc). Søjlerne ”Syg” og ”Rask” henviser henholdsvis til om personerne var smittede med Divoc eller ikke ej ved testningssituationen. Mens rækkerne ”Testet positiv” og ”Testet negativ” henviser til om personerne testede positiv eller negativ.

	Syg	**Rask**	**Sum**
Testet positiv	22	248	270
Testet negativ	3	2,227	2,230
Sum	25	2,475	2,500

a Hvad er sandsynligheden for at en tilfældig person fra denne stikprøveundersøgelse har en positiv test?

b Hvad er sandsynligheden for at en tilfældig person fra denne stikprøveundersøgelse både har en positiv test og er smittet med Divoc?

c Hvad er sandsynligheden for at en tilfældig person fra denne stikprøveundersøgelse har en positiv test, hvis personen er smittet med Divoc?

d Hvad er sandsynligheden for at en tilfældig person fra denne stikprøveundersøgelse har en positiv test, hvis personen ikke er smittet med Divoc? Sammenlign dette med spørgsmål c), hvad betyder det?

e Hvad er sandsynligheden for at en tilfældig person fra denne stikprøveundersøgelse er smittet med Divoc, hvis personens test er positiv? Hvad er sandsynligheden for at personen er rask, hvis personens test er positiv? Sammenlign dette. Hvad betyder det?

A.3 Second Exercise - The Second Contingency Table (Danish Version)

Tabellen viser opdigtet data fra en stikprøve der skal hjælpe med at afklare effekten af at masseteste en befolkning for en bestem opdigtet sygdom (kunne hedde Divoc). Søjlerne ”Syg” og ”Rask” henviser henholdsvis til om personerne var smittede med Divoc eller ikke ej ved testningssituationen. Mens rækkerne ”Testet positiv” og ”Testet negativ” henviser til om personerne testede positiv eller negativ.

	Syg	**Rask**	**Sum**
Testet positiv	15	922	937
Testet negativ	25	1,538	1,563
Sum	40	2,460	2,500

Er denne test for Divoc en god test? Hvorfor/hvorfor ikke?

A.4 Third Exercise - The Urn Problem (Danish Version)

I en urne ligger der fire kugler. To af dem er hvide og to af dem er sorte. Du trækker en kugle fra urnen og ser at kuglen er hvid, du putter kuglen i lommen.

a) Hvad er sandsynligheden for, at den næste kugle, du trækker, også er hvid?

Du putter alle kuglerne tilbage i urnen og starter eksperimentet forfra. Denne gang trækker du en kugle, men kigger ikke på kuglen inden du putter den i lommen. Du trækker en ny kugle, som du ser, er hvid.

b) Hvad er sandsynligheden for, at den kugle, der ligger i din lomme, også er hvid?

c) Er hændelserne, første kugle er hvid og anden kugle er hvid, afhængige eller uafhængige?

Tillægsopgave:

Du har den samme urne med de samme fire kugler i. Denne gang trækker du en kugle, og du ser den er hvid. Derefter putter du kuglen tilbage i urnen.

d) Hvad er sandsynligheden for, at den næste kugle, du trækker, også er hvid?

Gentag eksperimentet, hvor du trækker en kugle, som du viser til din ven, uden at du ser den, hvorefter du putter kuglen til bare i urnen. Nu trækker du en ny kugle, der viser dig at være hvid.

e) Hvad er sandsynligheden for, at din ven så en hvid kugle?

f) Er hændelserne, første kugle er hvid og anden kugle er hvid, denne gang afhængige eller uafhængige?

A.5 Fourth Exercise - The Monty Hall Problem (Danish Version)

Antag, at du deltager i et gameshow og får valget mellem at vælge en af tre døre. Bag den ene dør er der en bil og bag hver af de to andre døre er der en ged. Når du har valgt en dør, lad os sige nr. 1 (som er lukket), åbner værten, som ved, hvad der er bag hver dør, en anden dør (f.eks. nr. 3), som indeholder en ged. Du får nu mulighed for at ændre dit valg til dør nr. 2 eller holde dig til dør nr. 1. Hvad vil du gøre?

A.6 Fifth Exercise - The Drawer Problem (Danish Version)

Du har tre kommoder med netop to skuffer i hver. I en af kommoderne ligger der en sølvmønt i hver af skufferne. I en anden af kommoderne ligger der en guldmønt i hver af skufferne. I den sidste kommode ligger der en sølvmønt i den ene skuffe og en guldmønt i den anden. Nu vælger du en tilfældig skuffe i en tilfældig kommode.

a) Hvad er sandsynligheden for, at der er en guldmønt i den skuffe du har valgt?

Nu åbner du skuffen og ser, at der rent faktisk er en guldmønt i skuffen.

b) Hvad er sandsynligheden for, at der også er en guldmønt i den anden skuffe i kommoden?

I eksemplet fra første del, er sandsynligheden for, at der er en guldmønt i den skuffe du vælger først, og for, at der er en guldmønt i den anden skuffe i kommoden, du har valgt en skuffe i, uafhængige?

A.7 Sixth Exercise - The Three-Event Problem (Danish Version)

Du har to mønter, én 1-krone og én 2-krone. Du vil slå ”plat eller krone” med begge mønter samtidig, men vil først overveje sandsynlighederne for følgende tre hændelser:

- 1-kronen bliver krone.
- 2-kronen bliver krone.
- Netop 1 af mønterne bliver krone. Det vil sige hverken flere eller færre.

Efter disse indledende overvejelser, skal du overveje om alle tre hændelser er uafhængige af hinanden

A.8 The Initial Comprehension Test

You must now answer a few questions, which will be used as part of an empirical basis for the study you will participate in this week. This means that this test will not be handed over to your teacher or other employees at Stenhus Gymnasium and HF. If extracts from your answers are used in the thesis, these extracts will appear completely anonymous, where neither school, class nor names will be identifiable. All questions are about probability theory.

1. Describe in your own words what independence means in probability theory.
2. How do you find out whether two events (e.g. if a random person from your class is born in March and a random person from your class has an iPhone) are independent or not?
3. Determine whether the following events are independent:

    a. The events that a random car is black and the event that the driver of a random car is a woman, when you know that 10% of cars in Denmark are black and that 10 of cars driven by women are also black.

    b. The events that a random 18-year-old lives in Holbæk and that a random 18-year-old has a driving license, when you know that the probability that an 18-year-old lives in Holbæk is 5%, the probability that an 18-year-old has a driving license is 60 % and the probability that an 18-year-old both lives in Holbæk and has a driving license is 4%.

    c. The events that it rains on a random day and the event that a random day is a Monday, when you know that the probability that it rains is 20% and the probability that it rains on a Monday is 25%.

A.9 The Final Comprehension Test

You must now answer a few questions, which will be used as part of an empirical basis for the study you will participate in this week. This means that this test will not be handed over to your teacher or other employees at Stenhus Gymnasium and HF. If extracts from your answers are used in the thesis, these extracts will appear completely anonymous, where neither school, class nor names will be identifiable. All questions are about probability theory.

1. Describe in your own words what independence means in probability theory.
2. How do you find out whether two events (e.g. if a random person from your class is born in March and a random person from your class has an iPhone) are independent or not?
3. Determine whether the following events are independent:
   a. The events that a random 10-year-old swim as a hobby and that a random 10-year-old is a boy when you know that 14% of 10-year-olds swim as a hobby and that 14% of 10-year-old boys swim as a hobby.
   b. The events that a random person has passed mathematics at A-level and that a random person is among the richest half of Denmark's population, when you know that the probability that a random person has passed mathematics at A-level is 12%, the probability that a random person is among the richest half in Denmark is 50% and the probability that a random person has both passed mathematics at A-level and is among the richest half of the Danish population is 8%.
   c. The events that a random person has blue eyes and the event that a random person is born on a Monday, when you know that the probability of being born on a Monday is $\frac{1}{7}$ and the probability that a random person was born on a Monday if the person has blue eyes is $\frac{3}{14}$

A.10 The Initial Comprehension Test (Danish Version)

Du skal nu svare på et par spørgsmål, som skal bruges som en del af et empirisk grundlag for den undersøgelse I skal medvirke til i denne uge. Dermed bliver denne undersøgelse ikke udleveret til jeres lærer eller andre ansatte på Stenhus Gymnasium og HF. Hvis der bliver brugt uddrag fra jeres besvarelser i den færdige opgave, vil disse uddrag forekomme fuldstændigt anonymt, hvor hverken skole, klasse eller navne vil kunne identificeres. Alle spørgsmål handler om sandsynlighedsregning.

1. Beskriv med dine egne ord, hvad uafhængighed betyder indenfor sandsynlighedsregning.

2. Hvordan finder man ud af om to hændelser (f.eks. om en tilfældig fra jeres klasse er født i marts og en tilfældig fra jeres klasse har en iPhone) er uafhængige eller ej?

3. Afgør om følgende hændelser er uafhængige:

   a. Hændelserne at en tilfældig bil er sort og hændelsen at en tilfældig bils chauffør er en kvinde, når du ved at 10% af biler i Danmark er sorte og at 10% af biler kørt af kvinder også er sorte.

   b. Hændelserne at en tilfældig 18-årig bor i Holbæk og at en tilfældig 18-årige har kørekort, når du ved at sandsynligheden for at en 18-årig bor i Holbæk er 5%, sandsynligheden for at en 18-årig har kørekort er 60% og sandsynligheden for at en 18-årig både bor i Holbæk og har kørekort er 4%.

   c. Hændelserne at det regner på en tilfældig dag og hændelsen at en tilfældig dag er en mandag, når du ved at sandsynligheden for at det regner er 20% og sandsynligheden for at det regner på en mandag er 25%.

A.11 The Final Comprehension Test (Danish Version)

Du skal nu svare på et par spørgsmål, som skal bruges som en del af et empirisk grundlag for den undersøgelse I skal medvirke til i denne uge. Dermed bliver denne undersøgelse ikke udleveret til jeres lærer eller andre ansatte på Stenhus Gymnasium og HF. Hvis der bliver brugt uddrag fra jeres besvarelser i den færdige opgave, vil disse uddrag forekomme fuldstændigt anonymt, hvor hverken skole, klasse eller navne vil kunne identificeres. Alle spørgsmål handler om sandsynlighedsregning.

1. Beskriv med dine egne ord, hvad uafhængighed betyder indenfor sandsynlighedsregning.

2. Hvordan finder man ud af om to hændelser (f.eks. om en tilfældig fra jeres klasse er født i marts og en tilfældig fra jeres klasse har en iPhone) er uafhængige eller ej?

3. Afgør om følgende hændelser er uafhængige:

   a. Hændelserne at en tilfældig 10-årig går til svømning og at en tilfældig 10-årig er en dreng, når du ved at 14% af 10-årige går til svømning og at 14% af 10-årige drenge går til svømning.

   b. Hændelserne at en tilfældig har bestået matematik på A-niveau og at en tilfældig er blandt den rigeste halvdel af Danmarks befolkning, når du ved at sandsynligheden for at en tilfældig person har bestået matematik på A-niveau er 12%, sandsynligheden for at en tilfældig person er blandt den rigeste halvdel i Danmark er 50% og sandsynligheden for at en tilfældig både har bestået matematik på A-niveau og er blandt den rigeste halvdel af den danske befolkning er 8%.

   c. Hændelserne at en tilfældig person har blå øjne og hændelsen at en tilfældig person er født på en mandag, når du ved at sandsynligheden for at være født på en mandag er $\frac{1}{7}$ og sandsynligheden for at en tilfældig person er født på en mandag, hvis personen har blå øjne, er $\frac{3}{14}$